

STIC-Biotech/ChemLib

169186

From: Swope, Sheridan
Sent: Wednesday, September 28, 2005 3:41 PM
To: STIC-Biotech/ChemLib
Subject: 09/858332

For 09/858332, pls search:

SID 15: against the NT and AA data bases.

Collect the top 100 hits.

Thanks,

Sheridan Swope, Ph.D.
Patent Examiner, AU 1656
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

Barbo Bryen

Need more hits

P2n.rge

P2n.rst

RECEIVED
SEP 28 2005
STIC/CHL BLD
(STIC)

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIS: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2005, 11:22:37 ; Search time 118 Seconds
(without alignments)
39.332 Million cell updates/sec

Title: US-09-858-332G-15
Perfect score: 84
Sequence: 1 HNNNNNNNNNN 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 100 summaries

Database :

A_Geneseq_16Dec04:*
1: geneseq1980s:*
2: geneseq1980s:*
3: geneseq2000s:*
4: geneseq2001s:*
5: geneseq2002s:*
6: geneseq2003as:*
7: geneseq2003bs:*
8: geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	100.0	12	6	ABU08447 Affinity
2	84	100.0	13	6	ABP55170 Hexa-hist
3	84	100.0	16	2	AAW30240 HNI16, amp
4	84	100.0	16	2	AAV32981 Membrane
5	84	100.0	16	7	AAE38878 Membrane
6	84	100.0	16	8	AD004524 Potential
7	84	100.0	16	8	AD004531 Potential
8	84	100.0	20	6	ABU08460 Affinity
9	84	100.0	55	8	ADP32690 Receptor
10	84	100.0	55	8	ADP32690 Receptor
11	84	100.0	120	7	ADP59240 Rat Prote
12	84	100.0	120	7	ADP59240 Rat Prote
13	84	100.0	232	3	AAV32981 Membrane
14	69	82.1	1284	4	ABP31869 Drosophi
15	69	82.1	1284	4	ABP31869 Drosophi
16	68.5	81.5	697	4	ABP31869 Drosophi
17	67	79.8	105	5	ABP31869 Drosophi
18	66	78.6	16	2	AAW30247 HNI16, am
19	66	78.6	16	2	AAW30248 HNI16, am
20	66	78.6	16	2	AAV32981 Membrane
21	66	78.6	16	7	AAE38878 Membrane
22	66	78.6	16	7	AAE38878 Membrane
23	66	78.6	16	7	AAE38878 Membrane
24	66	78.6	16	8	AD004531 Potential
25	66	78.6	16	8	AD004532 Potential

26	66	78.6	16	8	ADP31869 Drosophi
27	66	78.6	16	8	ADP31869 Drosophi
28	65	77.4	124	4	ABP31869 Drosophi
29	65	77.4	103	5	ABP31869 Drosophi
30	64	76.2	193	3	AAE38878 Membrane
31	64	76.2	193	3	AAE38878 Membrane
32	64	76.2	244	3	AAE38878 Membrane
33	64	76.2	244	3	AAE38878 Membrane
34	64	76.2	244	3	AAE38878 Membrane
35	64	76.2	300	3	AAE38878 Membrane
36	64	76.2	301	3	AAE38878 Membrane
37	64	76.2	451	6	ABU08447 Affinity
38	64	76.2	913	3	AAE38878 Membrane
39	64	76.2	923	3	AAE38878 Membrane
40	64	76.2	993	3	AAE38878 Membrane
41	64	76.2	1176	8	ADP31869 Drosophi
42	63	75.0	341	2	AAE38878 Membrane
43	62	73.8	150	2	AAE38878 Membrane
44	62	73.8	150	2	AAE38878 Membrane
45	62	73.8	150	2	AAE38878 Membrane
46	62	73.8	150	2	AAE38878 Membrane
47	62	73.8	150	2	AAE38878 Membrane
48	62	73.8	150	2	AAE38878 Membrane
49	62	73.8	150	2	AAE38878 Membrane
50	62	73.8	150	2	AAE38878 Membrane
51	62	73.8	150	2	AAE38878 Membrane
52	62	73.8	150	2	AAE38878 Membrane
53	62	73.8	150	2	AAE38878 Membrane
54	62	73.8	150	2	AAE38878 Membrane
55	62	73.8	150	2	AAE38878 Membrane
56	62	73.8	150	2	AAE38878 Membrane
57	62	73.8	150	2	AAE38878 Membrane
58	62	73.8	150	2	AAE38878 Membrane
59	62	73.8	150	2	AAE38878 Membrane
60	62	73.8	150	2	AAE38878 Membrane
61	62	73.8	150	2	AAE38878 Membrane
62	62	73.8	150	2	AAE38878 Membrane
63	62	73.8	150	2	AAE38878 Membrane
64	62	73.8	150	2	AAE38878 Membrane
65	61	72.6	429	6	ABU08447 Affinity
66	61	72.6	850	4	ABP31869 Drosophi
67	61	72.6	1028	4	ABP31869 Drosophi
68	59	70.2	59	4	ABG30585 Novel hum
69	59	70.2	83	3	AAE38878 Membrane
70	59	70.2	120	3	AAE38878 Membrane
71	59	70.2	124	4	ABG30585 Novel hum
72	59	70.2	135	3	AAE38878 Membrane
73	59	70.2	171	5	ABG30585 Novel hum
74	59	70.2	186	8	ABG30585 Novel hum
75	59	70.2	281	6	ABG30585 Novel hum
76	59	70.2	292	4	AAE38878 Membrane
77	59	70.2	292	4	AAE38878 Membrane
78	59	70.2	292	4	AAE38878 Membrane
79	59	70.2	292	4	AAE38878 Membrane
80	59	70.2	292	4	AAE38878 Membrane
81	59	70.2	292	4	AAE38878 Membrane
82	59	70.2	292	4	AAE38878 Membrane
83	59	70.2	292	4	AAE38878 Membrane
84	59	70.2	292	4	AAE38878 Membrane
85	59	70.2	292	4	AAE38878 Membrane
86	59	70.2	292	4	AAE38878 Membrane
87	59	70.2	374	4	ABP31869 Drosophi
88	59	70.2	467	8	ADP31869 Drosophi
89	59	70.2	537	4	ABP31869 Drosophi
90	59	70.2	989	4	ABP31869 Drosophi
91	59	70.2	1300	4	ABP31869 Drosophi
92	58	69.0	19	2	AAW30247 HNI16, am
93	58	69.0	19	2	AAW30248 HNI16, am
94	58	69.0	117	4	ABG30585 Novel hum
95	58	69.0	156	4	ABG30585 Novel hum
96	58	69.0	156	4	ABG30585 Novel hum
97	58	69.0	156	4	ABG30585 Novel hum
98	58	69.0	156	4	ABG30585 Novel hum

Abg46175	Human	pep
Adc38684	Human	sec

ALIGNMENTS

ABU08447 standard; peptide; 12 AA.

DT 17-JUN-2003 (first entry)

DE Affinity purification peptide #2.
XX

KW metal ion affinity peptide; ion protein; protein purification;
KW metal chelating resin; metal ion chelate resin; hard metal; Fe3+;
KW Al3+; Co2+; Cu2+; Ni3+; Zn2+; Co2+; protein-protein interaction;
KW metal ion affinity peptide-tagged recombinant protein; IMC;
KW DNA-protein interaction; immobilised metal ion affinity chromatography;
KW gene expression; phosphorylation state.

Synthetic

PN US2002164718-A1

PD 07-NOV-2002

15-MAY-2001; 2001US-00858332.

PR 25-SEP-1998; 98US-0101867P.

PR 23-SEP-1999; 99US-00404017.
XX

PA (TCHA/) TCHAGA G S.
PA (JOKH/) JOKHADZE G G.

PI Tchaga GS, Jokhadze GG;

WPI: 2003-361747/34

XX
OK
N-PSDB; ABX94280

new metal ion affinity peptide useful, when fused to a fusion partner polypeptide, for protein purification methods and to study protein-protein interactions and nucleic acid-protein interactions.

Example 1; Fig 3; 23pp; English.

The present invention relates to metal ion affinity peptides, fusion proteins containing metal ion affinity peptides, and polynucleotide sequences encoding the fusion proteins. The presence of a metal ion affinity peptide in a fusion protein allows purification of the fusion protein on a metal chelating resin. The method involves contacting a sample comprising a fusion protein with a metal ion chelate resin comprising a first metal ion, preferably a hard metal ion such as Ca^{2+} or Al^{3+} and eluting any resultant bound fusion protein from the resin. The resin comprises an immobilised second metal ion, such as Fe^{3+} , which competes with the first metal ion for the fusion protein. The method comprises contacting the sample with a second immobilised metal ion and eluting any resultant bound fusion protein from the first and second resins. The second metal ion is an intermediate metal ion such as Cu^{2+} , Ni^{2+} , Zn^{2+} or Co^{2+} . The metal ion is an intermediate metal ion and eluting any resultant bound fusion protein from the first and second resins. The method is useful for the study of protein-protein interactions and nucleic acid molecule-protein interactions. Using solid phase immobilised metal ion affinity chromatography (IMAC). They are also useful in high throughput systems which find use in massive parallel gene expression experiments, e.g. to determine the effect of an agent on synthesis of a protein or set of proteins, to analyse developmental stage-specific, or tissue-specific synthesis of a protein and to analyse the phosphorylation state of a protein. These methods find use in applications to characterise a protein of unknown identity or function, and in enzymatic reactions. ABU08446-ABU08450 represent affinity purification peptides

XX	Sequence 12 AA;
SQ	

Query Match	100.0%;	Score 84;	DB 6;	Length 12;
Best Local Similarity	100.0%;	Pred. No. 6.5e-06;		
Matches 12; Conservative	0;	Mismatches 0;	Indels	

Qy	1 HNNHNNHHNHNEN 12
D_b	1 HNNHNNHHNHNHN 12

RESULT 2

ID	ABP55170	standard; peptide; 13 AA
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

AC ABP55170

11-FEB-2003 (Fri Feb 2003)

Hexahistidine-asparagine tag peptide.

Vector; splice donor; intron; affinity tag

XX

MOZ00283910-A2.
XX

PD 24-OCT-2002.
XX

17-JAN-2002; 2002WO-US001604

18-JAN-2001; 2001US-0263358P

(CLON-) CLONTECH LAB INC.

FARMEX AA

WPI; 2003-075553/07.

acceptor vector having at least one splice site, where each vector comprises at least one recombination site, useful in preparing an intron-containing vector.

Example 2; Page 42; 64pp; English.

The present sequence is the peptide sequence of a 6xHisAMN peptide tag, which can be encoded by vectors of the invention to facilitate protein purification in bacteria. The invention provides methods for producing a vector that includes at least one spliceable intron. The vector is produced from donor and acceptor vectors that each include a site-specific recombinase site, as well as splice donor and acceptor sites that, upon site-specific recombination of the donor and acceptor vectors, define an intron in the product vector of the recombination. Examples of donor, acceptor and product vectors are given in ABV75997-ABV76000. The method is useful for production of vectors encoding C-terminal tagged fusion proteins, and expression vectors encoding pure protein. Rapid transfer of a DNA molecule from one vector to another is achieved in vitro or in vivo without the need to rely upon restriction enzyme digestions.

Query Match:

Query Match	100.0%	Score 84;	DB 6;	length 13;
Best Local Similarity	100.0%	Pred. No. 7.1e-06;		
Matches 12;	Conservative 0;	Mismatches 0;	Indels 0;	Cuts 0;

1 HHHHHHHHHHNN 12
| | | | | | | | | |
2 HHHHHHHHHHNN 13

RESULT 3
 AAM30240
 ID AAM30240 standard; peptide; 16 AA.
 AC
 XX AAM30240;
 XX
 DT 25-MAR-2003 (revised)
 DT 21-JAN-1998 (first entry)
 XX
 DE HN16, amphiphilic potential membrane forming peptide.
 XX
 KW Membrane-forming peptide; amphiphilic; monovalent metal cation; suture;
 KW internal lining; slow-diffusion drug delivery vehicle; erythropoietin;
 KW tissue-type plasminogen activator; haemoglobin; insulin; artificial skin;
 KW separation matrix; dialysis membranes; filter; collagen; cell migration;
 KW Alzheimer's disease; scrapie infection; therapy; proteolytic digestion;
 KW acidic pH resistant; stomach acid; cell adhesion; cell monolayer;
 KW tissue culture; differentiated cell; stratified cell layer.
 XX
 OS Synthetic.
 OS
 PN US5670483-A.
 PN
 PD 23-SEP-1997.
 PD
 PE 30-NOV-1994; 94US-00346849.
 PE
 PR 28-DEC-1992; 92US-00973326.
 PR
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
 PA
 XX Rich A, Holmes T, Zhang S, Lockshin C;
 PI
 XX WPI, 1997-479506/44.
 XX
 DR
 PT bio:material(s), separation matrices, drug delivery vehicles, etc.
 PT
 PS Claim 30; Col 64; 49pp; English.
 PS
 CC AAM30219-W30252 are amphiphilic peptides used in the macroscopic membrane
 CC (MM) of the invention. The MM is formed by self-assembly of the
 CC amphiphilic peptides (with alternating hydrophilic and hydrophobic amino
 CC acids) in an aqueous medium containing monovalent metal cations. The MM
 CC are stable in serum, are non-cytotoxic, and are useful in biomaterial
 CC applications, such as medical products (e.g. sutures), or internal
 CC lining. The MM are useful as slow-diffusion drug delivery vehicles for
 CC protein-type drugs, including erythropoietin, tissue-type plasminogen
 CC activator, synthetic haemoglobin and insulin. They can be used in
 CC applications, such as separation matrices (e.g., dialysis membranes). The
 CC extremely small pore size (interfilament distance) of the MM makes them
 CC useful as filters. Collagen may be combined with the peptides to produce
 CC membranes for use as artificial skin. The MM may be used for making very
 CC thin, transparent fabric. Drugs which inhibit the self assembly of the
 CC peptides into filaments or filamentous membranes may be useful for
 CC treating Alzheimer's disease or scrapie infection. As they are resistant
 CC to proteolytic digestion and alkaline and acidic pH (such as stomach
 CC acid), drug delivery vehicles made of the MM could be taken orally. The
 CC charged residues and conformation of the MM promote cell adhesion and
 CC migration. The permeability of the MM also permits diffusion of small
 CC molecules, to the underside of cell monolayers, useful for tissue culture
 CC of differentiated cells and/or stratified cell layers. (Updated on 25-MAR-
 CC -2003 to correct PF field.)
 CC
 SQ Sequence 16 AA;
 XX

Db 1 HNHHNNHNHNHN 12

RESULT 4
ID AAY32981
AC AAY32981 standard; peptide; 16 AA.
XX AAY32981;
XX
XX
XX
XX 09-NOV-1999 (first entry)
XX
XX
XX Membrane forming aphiphilic peptide HN16.
XX
XX Membrane forming peptide; cell culturing; macroscopic membrane;
XX amphiphilic peptide; slow-diffusion drug delivery system; cell growth;
XX artificial skin; separation matrix; artificial tissue; scrapie infection;
XX Alzheimer's disease; liver cirrhosis; kidney amyloidosis;
XX protein conformational disease.
XX
XX Synthetic.
XX
XX US5955343-A.
XX
XX 21-SEP-1999.
XX
XX 22-AUG-1994; 94US-00293284.
XX
XX 28-DEC-1992; 92US-00973326.
XX
XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Rich A, Dipersio CM, Lockshin C, Holmes T, Zhang S;
XX
XX WPI; 1999-539576/45.
XX
XX Cell cultures utilizing stable macroscopic membranes formed by the self-
XX assembly of amphiphilic peptides.
XX
XX Example 5; Col 27; 49pp; English.
XX
XX This sequence represents a membrane forming aphiphilic peptide used in
XX the in vitro method of the invention. The method is for culturing cells
XX utilizing stable macroscopic membranes formed by the self-assembly of
XX amphiphilic peptides. The peptides are salt-induced to form insoluble and
XX protease-resistant protein filaments with a beta-sheet secondary
XX structure. The membranes may be may be useful in a wide variety of
XX medical, research, industrial and biomaterial applications such as slow-
XX diffusion drug delivery systems, artificial skin and separation matrices.
XX The membranes may be used to support in vitro cell attachment and growth
XX and for supporting artificial tissue (e.g. for in vivo use as implants).
XX They are particularly useful as experimental models for Alzheimer's
XX disease and scrapie infection and so may be used in disease modelling
XX experiments and to assay for agents which modulate the disease processes.
XX Additionally, they may be used in this way to study liver cirrhosis,
XX kidney amyloidosis and other protein conformational diseases. The
XX membrane in the membrane/cell mixture produced in the method: (i)
XX supports cellular attachment and growth; (ii) self-assembles to form
XX large, macroscopic membranes that are insoluble and stable in aqueous
XX solutions, serum and ethanol; (iii) is highly resistant to heat
XX alkaline/acidic pHs, chemical denaturants and proteolytic digestion. (iv)
XX is non-cytotoxic and non-immunogenic; (v) is visible to the naked eye if
XX dyed but is otherwise transparent; (vi) may form thin, permeable, high
XX density sheets or fibril like structures with simple structures, high
XX tensile strength and a porous structure; (vii) may be metabolised by
XX humans and animals; (viii) is inexpensive to produce; and (ix) can be
XX produced and stored in sterile conditions

Sequence 16 AA;

```

Query Match      100.0%;  Score 84;  DB 2;  Length 16;
Best Local Similarity 100.0%;  Pred. No. 8.9e-06;
Matches 12;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

OY      1 HHHHHHHHHN 12
      |||||

```

```

Query Match      100.0%; Score 84; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.9e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 HNNNNNNNNNN 12
Db 1 HNNNNNNNNNN 12

RESULT 5

ID AAE38878 standard; peptide; 16 AA.

AAE38878;

18-DEC-2003 (first entry)

Membrane forming peptide, HNI6.

Macroscopic membrane; biomaterial; medical product; internal lining; artificial skin; slow-diffusion drug delivery; chromatography column; separation matrix; suture; dialysis membrane; viral filter; fabric; pathology; Alzheimer's disease; scrapie infection.

Unidentified.

US6548630-B1.

15-APR-2003.

22-JUL-1997; 97US-00898300.

28-DEC-1992; 92US-00973326.

30-NOV-1994; 94US-00346849.

(MASI) MASSACHUSETTS INST TECHNOLOGY.

Zhang S, Lockshin C, Rich A, Holmes T;

WPI; 2003-584339/55.

Macroscopic membrane useful in biomaterial applications e.g. sutures is formed by self-assembly of amphiphilic peptides in an aqueous solution containing monovalent metal cations.

Example 5; Col 25; 51pp; English.

The invention relates to a macroscopic membrane formed by self-assembly of amphiphilic peptides in an aqueous solution containing monovalent metal cations. The peptides have alternating hydrophobic and hydrophilic amino acids and are complementary and structurally compatible. The membranes are useful as biomaterial for medical products (e.g. sutures, artificial skin, internal linings), as vehicles for slow-diffusion drug delivery (preferably for protein type drugs e.g. erythropoietin, tissue separation matrices (e.g. dialysis membranes, chromatography columns, filters for the removal of viruses and other microscopic contaminants), for other uses requiring permeable and water-insoluble material (e.g. for culturing cell monolayers including differentiated cells and/or stratified cell layers), for the preparation of very thin and transparent fabric and as a model system for investigating the properties of biological protein structures and providing insights into the pathology and potential treatment of conditions involving the presence of the proteins and proteinaceous structures (e.g. Alzheimer's disease, scrapie infection and in origin of life studies related to cell membranes and cellular compartmentalisation). The present sequence is a membrane forming peptide used in the invention

Sequence 16 AA;

Query Match

Best Local Similarity 100.0%; Score 84; DB 7; Length 16;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12

Db 1 HNNNNNNNNNN 12

RESULT 6

ID ADO04524 standard; peptide; 16 AA.

ADO04524;

29-JUL-2004 (first entry)

Potential membrane-forming peptide, HNI6.

Macroscopic membrane; cell growth; EAK16; Z-DNA binding protein; zoetin; medical product; suture; artificial skin; internal lining; slow-diffusion drug delivery system; protein-type drug; erythropoietin; haemoglobin; insulin; cell adhesion; cell migration; Alzheimer's disease; scrapie infection.

Unidentified.

US2004087013-A1.

06-MAY-2004.

17-MAR-2003; 2003US-00390472.

28-DEC-1992; 92US-00973326.

22-AUG-1994; 94US-00293284.

26-MAR-1997; 97US-00824515.

(HOLM/) HOLMES T.

(ZHAN/) ZHANG S.

(RICH/) RICH A.

(DIPE/) DIPERSIO C M.

(LOCK/) LOCKSHIN C.

Holmes T, Zhang S, Rich A, Dipersio CM, Lockshin C;

WPI; 2004-356208/33.

Novel EAK16 protein incorporated into macroscopic membranes, useful in biomaterial applications such as medical products, artificial skin or internal linings, slow-diffusion drug delivery systems for in vitro cell growth.

Example 5; SEQ ID NO 53; 56pp; English.

The invention relates to a method for in vitro cell culture which involves adding a macroscopic membrane that is formed by self-assembly of amphiphilic peptide in an aqueous solution containing monovalent metal cations to a cell culture medium comprising cells, thus forming a membrane/culture mixture, maintaining the mixture under conditions sufficient for cell growth. The invention also relates to EAK16 peptide derived from yeast Z-DNA binding protein (zoetin). Zoetin incorporated into the macroscopic membranes are useful in biomaterial applications such as medical products (e.g. sutures), artificial skin or internal linings, slow-diffusion drug delivery systems supports for in vitro cell growth or culture and support for artificial tissue for in vivo use, as e.g. erythropoietin, synthetic haemoglobin, insulin, etc., useful as slow-diffusion drug delivery vehicle for delivering protein-type drugs CC conductive biopolymer for culturing cell monolayers, for promoting cell CC adhesion and migration, useful as experimental models for Alzheimer's CC disease and scrapie infection. The present sequence is a potential CC membrane-forming peptide used to illustrate the method of the invention.

Sequence 16 AA;

Query Match

Best Local Similarity 100.0%; Score 84; DB 8; Length 16;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12

|||||

CC e.g. to determine the effect of an agent on synthesis of a protein or set
CC of proteins, to analyse developmental stage-specific, or tissue-specific
CC synthesis of a protein and to analyse the phosphorylation state of a
CC protein. These methods find use in applications to characterise a protein
CC of unknown identity or function, and in enzymatic reactions. The present
CC sequence represents an affinity peptide
SQ Sequence 20 AA;

Query Match 100.0%; Score 84; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
Db 1 HNNHNNHNNHNN 12

RESULT 9

ADP32690
ID ADP32690 standard; protein; 55 AA.

AC ADF32690;

DT 26-FEB-2004 (first entry)

DE Receptor detection related peptide SEQ ID NO:12.

KW binding event; protein receptor; beta-galactosidase; detection;
KM enzyme active site.

OS Synthetic.

PN WO2003102154-A2.

PD 11-DEC-2003.

PF 29-MAY-2003; 2003WO-US017428.

PR 29-MAY-2002; 2002US-0384060P.

PA (DISC-) DISCOVERX INC.

PI Nagvi T, Rouhani R, Singh R;

DR WPI; 2004-043100/04.

PT Performing a determination of a binding event of a protein receptor (e.g.
PT enzyme) with a ligand (e.g. a steroid or a nucleotide) comprises
PT employing a beta-galactosidase fragment as an enzyme donor.

PS Example; SEQ ID NO 12; 38pp; English.

XX The present invention describes a method for performing a determination
CC of a binding event of a protein receptor with a ligand, where the protein
CC receptor is other than an antibody or its polyvalent fragment, and where
CC enzyme donor fragment (ED) of beta-galactosidase is used as the label for
CC complexation with enzyme acceptor fragment (EA), the improvement
CC comprising employing as the enzyme donor a beta-galactosidase fragment of
CC 36-50 amino acids. Also described: (1) detecting the active site of an
CC enzyme, comprising combining in an assay medium a complex formed from a
CC sample suspected of comprising the enzyme and an ED conjugate of ED of 36
CC amino acids bonded to an enzyme binding site inhibitor having a
CC binding constant of at least about 10⁻⁸ M, with EA and a beta-
CC galactosidase substrate producing a detectable product; and detecting the
CC resulting signal as indicative of the presence of the enzyme; (2) a
CC complex of a receptor and a conjugate of a ligand binding the receptor
CC bonded to an ED of 36-50 amino acids; and (3) a kit comprising an enzyme
CC donor, a beta-galactosidase fragment of 36-50 amino acids conjugated to a
CC receptor ligand, a source of a beta-galactosidase enzyme acceptor,
CC optionally a beta-galactosidase substrate, and instructions for
CC performing the method cited above. The method is useful in detecting
CC protein receptors using labelled ligands. The method can also be used in

CC developing reagents that permit sensitive detection of receptors, that
CC allow for competitive assays and that may be readily produced as
CC conjugates or fused proteins. The kit may be used in marketing the
CC product and in research and commercial settings. The present sequence
CC represents an amino acid sequence which is used in the exemplification of
CC the present invention.
SQ Sequence 55 AA;

Query Match 100.0%; Score 84; DB 8; Length 55;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
Db 1 HNNHNNHNNHNN 12

RESULT 10

ADJ55662
ID ADJ55662 standard; peptide; 55 AA.

AC ADJ55662;

DT 06-MAY-2004 (first entry)

DE ED peptide of beta-galactosidase with purification & cleavage tag Seq 16.

KW enzyme donor; beta-galactosidase; competitive assay;
KM non-competitive assay; intracellular event; non-interfering label; ED.

OS Synthetic.

PN WO2003093786-A2.

PD 13-NOV-2003.

PF 24-APR-2003; 2003WO-US012589.

PR 02-MAY-2002; 2002US-0376935P.

PA (DISC-) DISCOVERX INC.

PI Nagvi T, Rouhani R, Singh R;

DR WPI; 2004-011925/01.

PT New enzyme donor for complementation with an enzyme acceptor fragment of
PT beta-galactosidase, useful for performing competitive and non-competitive
PT assays or for monitoring intracellular events.

PS Disclosure; SEQ ID NO 16; 29pp; English.

XX This invention relates to novel peptides that provide short enzyme donor
CC (ED) fragments capable of forming a complex with a large fragment of beta
CC -galactosidase (enzyme acceptor) to form a functional enzyme.
CC Specifically, it refers to enzyme donor fragments that substantially
CC constitute the natural sequence of the N-terminal proximal sequence of
CC beta-galactosidase, and as such can provide improved properties and
CC preparation due to their reduced size. The present invention describes
CC these short oligopeptide enzyme donors as labels that can linked to a
CC variety of compounds of interest, particularly proteins, in order to
CC determine the status of the compound, serving as a mimic of the natural
CC compound and for isolation purposes e.g. protease recognition sequences.
CC Furthermore, the enzyme donor is useful for performing competitive and
CC non-competitive assays, monitoring intracellular events or other
CC processes where a sensitive non-interfering label is desired. This
CC peptide sequence is an enzyme donor peptide fragment containing an N-
CC terminal 6(His-Asn) tag with a C-terminal thrombin cleavage site, given
CC in an exemplification of the invention.
SQ Sequence 55 AA;

Query Match 100.0%; Score 84; DB 8; Length 55;
 Best Local Similarity 100.0%; Pred. No. 3.3e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
 |||||
 DB 1 HNNHNNHNNHNN 12

RESULT 11

ADE59240
 ID ADE59240 standard; protein; 120 AA.

AC ADE59240;

DT 29-JAN-2004 (first entry)

DE Rat Protein AAF75130, SEQ ID NO 5133.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002MO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

DR WPI; 2003-268312/26.

DR GENBANK; AAF75130.

PT New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at

CC ftp://wipo.int/pub/published_pct_sequences.
 XX Sequence 120 AA;

QY 1 HNNHNNHNNHNN 12
 |||||
 DB 29 HNNHNNHNNHNN 40

Query Match 100.0%; Score 84; DB 7; Length 120;
 Best Local Similarity 100.0%; Pred. No. 7.7e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 12

ADE59237
 ID ADE59237 standard; protein; 120 AA.

AC ADE59237;

DT 29-JAN-2004 (first entry)

DE Rat Protein AAF75130, SEQ ID NO 5129.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002MO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

DR WPI; 2003-268312/26.

DR GENBANK; AAF75130.

PT New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene

CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 120 AA;

Query Match 100.0%; Score 84; DB 7; Length 120;
 Best Local Similarity 100.0%; Pred. No. 7.7e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
 |||||
 DB 29 HNNHNNHNNHNN 40

RESULT 13
 AAY95033

ID AAY95033 standard; protein; 232 AA.

XX AAY95033;

DT 23-JUN-2000 (first entry)

XX Candida albicans polypeptide sequence # 1.

XX Candida albicans infection; growth; survival; medicament; AIDS;
 XX vulvovaginitis; immunocompromised patient; treat.

OS Candida albicans.

PN EP982401-A2.

XX 01-MAR-2000.

PF 23-DEC-1998; 98EP-00310694.

XX 14-AUG-1998; 98GB-00017796.

XX (JANNC) JANSSEN PHARM NV.

XX Contreras RH, Nellesen B, De Backer MD, Luyten WHML, Viaene JZ;
 PI Logshe MG;

XX WPI; 2000-258614/23.

XX Essential polypeptides isolated from Candida albicans, useful in the
 PT treatment of diseases caused by C. albicans, especially in
 PT immunocompromised subjects, e.g., AIDS patients.

XX Claim 3; Page 13; 133pp; English.

XX This sequence represents a polypeptide that is critical for the survival
 CC and growth of Candida albicans. The C. albicans nucleic acid molecules
 CC encoding the polypeptides of the invention may be used as probes and
 CC primers for detecting homologous nucleic acid molecule sequences. The
 CC polypeptides and nucleic acid molecules and compounds identified as
 CC selectively modulating the expression of the polypeptides, may be used as
 CC medicaments or for the preparation of a medicament to treat C. albicans
 CC associated diseases, especially in AIDS patients and to treat
 CC vulvovaginitis in otherwise healthy females. The use of the polypeptides
 CC and polynucleotide sequences to treat C. albicans associated diseases has
 CC fewer side effects and less toxicity than previously used methods such as
 CC the use of amphotericin. This method is therefore especially suitable for
 CC immunocompromised patients, such as AIDS patients

XX Sequence 232 AA;

Query Match 83.3%; Score 70; DB 3; Length 232;
 Best Local Similarity 83.3%; Pred. No. 0.014;
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
 |||||
 DB 7 HNNHNNHNNHNN 18

RESULT 14
 ABB71869

ID ABB71869 standard; protein; 1284 AA.

XX ABB71869;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 42399.

XX Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

XX 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.

XX N-PSDs; ABL15972.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.

XX Disclosure; SEQ ID NO 42399; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL16176-ABL30511) and the encoded proteins (ABBS7737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1284 AA;

Query Match 82.1%; Score 69; DB 4; Length 1284;
 Best Local Similarity 75.0%; Pred. No. 0.13;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
 |||||
 DB 1163 HHHNNHSHNNHNN 1174

RESULT 15
 ABB08376

ID ABB08376 standard; protein; 1284 AA.

XX ABB08376;

DT 07-MAY-2002 (first entry)

XX D. melanogaster CG10231 gene product.

XX transketolase-like; enzyme; fruitfly; cytosolic; antianaemic; cancer;
 KW anaemia; end-stage renal disease; sensory neuropathy; uraemia.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200192310-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-EP006125.
 XX
 PR 31-MAY-2000; 2000US-0207950P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Ramakrishnan S;
 XX
 DR WPI; 2002-130523/17.
 XX
 PT Novel isolated human transketolase-like enzyme polypeptide, useful for
 PT treating cancer, anemia, end-stage renal disease and sensory neuropathy
 PT associated with uremia.
 XX
 PS Disclosure; Fig 5; 91pp; English.
 XX
 CC The invention relates to an isolated transketolase-like enzyme
 CC polypeptide. The activity of the polypeptide of the invention may be
 CC described as cytosolic and antianaemic. Polypeptides and polynucleotides
 CC of the invention can be used to identify agents that modulate the
 CC activity of a transketolase-like enzyme. They are also useful for
 CC treating transketolase-like enzyme disorder such as cancer, anaemia, end-
 CC stage renal disease and sensory neuropathy associated with uraemia.
 CC Therapeutic agents of the invention are useful for modulating the
 CC activity of transketolase-like enzyme in the above mentioned diseases.
 CC Polypeptides and polynucleotides of the invention are useful for
 CC identifying test compounds which act as agonists or antagonists, for
 CC raising specific antibodies, and as a bait protein in a two-hybrid or
 CC three-hybrid assay. They are also useful in diagnostic assays for
 CC detecting diseases and abnormalities or susceptibility to disease and
 CC abnormalities related to the presence of mutations in the transketolase-
 CC like enzyme encoding sequence. Fusion proteins including the polypeptide
 CC of the invention may be useful in the generation of antibodies and in
 CC various assay systems. The current sequence represents the D.
 CC melanogaster CG10231 gene product that was used to aid the identification
 CC of human transketolase-like enzyme
 XX
 SQ Sequence 1284 AA;
 XX
 Query Match 82.1%; Score 69; DB 5; Length 1284;
 Best Local Similarity 75.0%; Pred. No. 0.13;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HNNHNNHNNHNN 12
 Db 1163 HHHNHSHNNHN 1174
 XX
 RESULT 16
 ID ABB64801 standard; protein; 697 AA.
 XX
 AC ABB64801;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 21195.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX

PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EM;
 XX
 DR WPI; 2001-656860/75.
 XX
 DR N-PSDB; ABL08904.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 21195; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 697 AA;
 XX
 Query Match 81.5%; Score 68.5; DB 4; Length 697;
 Best Local Similarity 84.6%; Pred. No. 0.077;
 Matches 11; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
 QY 1 HNNH-NHNNHNN 12
 Db 622 HNNHNNHNNHNS 634
 XX
 RESULT 17
 ID ABB27096 standard; protein; 105 AA.
 XX
 AC ABB27096;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 3368.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus agalactiae.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB004789.
 XX
 PR 27-OCT-2000; 2000GB-00026333.
 XX
 PR 24-NOV-2000; 2000GB-00028727.
 XX
 PR 07-MAR-2001; 2001GB-00005640.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PA (GENO-) INST GENOMIC RES.
 XX

PI Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C,
 PI Tettein H,
 DR WPI; 2002-352536/38.
 DR N-PSDB; ABN67727.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.
 XX
 PS Claim 1; Page 3490; 4525pp; English.
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and anti-inflammatory
 CC activity. (1), nucleic acids encoding (1), ABN6044-ABV71526 and
 CC antibodies that bind (1) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a
 CC biological sample. (1) is used to determine whether a compound binds to
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (1) may be used to recombinantly produce (1) and may be
 CC used in gene therapy. Antibodies to (1) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX
 SO Sequence 105 AA;
 Query Match 79.8%; Score 67; DB 5; Length 105;
 Best Local Similarity 81.8%; Pred. No. 0.016;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 HNNHNNHNNH 11
 Db 3 HNNHNDHNEH 13
 RESULT 18
 AAW30247 standard; peptide: 16 AA.
 ID AAW30247;
 AC AAW30247;
 XX
 DT 25-MAR-2003 (revised)
 DT 21-JAN-1998 (first entry)
 XX
 DE HNQ16, amphiphilic potential membrane forming peptide.
 XX
 KW Membrane-forming peptide; amphiphilic; monovalent metal cation; suture;
 KW internal lining; slow-diffusion drug delivery vehicle; erythropoietin;
 KW tissue-type plasminogen activator; haemoglobin; insulin; artificial skin;
 KW separation matrix; dialysis membranes; filter; collagen; cell migration;
 KW Alzheimer's disease; scrapie infection; therapy; proteolytic digestion;
 KW acidic pH resistant; stomach acid; cell adhesion; cell monolayer;
 KW tissue culture; differentiated cell; stratified cell layer.
 XX
 OS Synthetic.
 XX
 PN US5670483-A.
 PD 23-SEP-1997.
 XX
 PF 30-NOV-1994; 94US-00346849.
 XX
 PR 28-DEC-1992; 92US-00973326.
 XX
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX Rich A, Holmes T, Zhang S, Lockshin C;
 PI

XX
 DR WPI; 1997-479506/44.
 XX
 PT Membranes formed by self-assembly of amphiphilic peptide(s) - useful as
 PT bio:material(s), separation matrices, drug delivery vehicles, etc.
 XX
 PS Claim 30; Col 64; 49pp; English.
 CC AAW30219-W30252 are amphiphilic peptides used in the macroscopic membrane
 CC (MM) of the invention. The MM is formed by self-assembly of the
 CC amphiphilic peptides (with alternating hydrophilic and hydrophobic amino
 CC acids) in an aqueous medium containing monovalent metal cations. The MM
 CC are stable in serum, are non-cytotoxic, and are useful in biomaterial
 CC applications, such as medical products (e.g. sutures), or internal
 CC linings. The MM are useful as slow-diffusion drug delivery vehicles for
 CC protein-type drugs, including erythropoietin, tissue-type plasminogen
 CC activator, synthetic haemoglobin and insulin. They can be used in
 CC applications, such as separation matrices (e.g., dialysis membranes). The
 CC extremely small pore size (interfilament distance) of the MM makes them
 CC useful as filters. Collagen may be combined with the MM makes them
 CC membranes for use as artificial skin. The MM may be used for making very
 CC thin, transparent fabric. Drugs which inhibit the self assembly of the
 CC peptides into filaments or filamentous membranes may be useful for
 CC treating Alzheimer's disease or scrapie infection. As they are resistant
 CC to proteolytic digestion and alkaline and acidic pH (such as stomach
 CC acid), drug delivery vehicles made of the MM could be taken orally. The
 CC charged residues and conformation of the MM promote cell adhesion and
 CC migration. The permeability of the MM also permits diffusion of small
 CC molecules, to the underside of cell monolayers, useful for tissue culture
 CC of differentiated cells and/or stratified cell layers. (Updated on 25-MAR
 CC -2003 to correct PF field.)
 XX
 SO Sequence 16 AA;
 Query Match 78.6%; Score 66; DB 2; Length 16;
 Best Local Similarity 81.8%; Pred. No. 0.003;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 HNNHNNHNNH 11
 Db 1 HNNHNNHNNH 11
 RESULT 19
 AAW30248 standard; peptide: 16 AA.
 ID AAW30248;
 AC AAW30248;
 XX
 DT 25-MAR-2003 (revised)
 DT 21-JAN-1998 (first entry)
 XX
 DE HNQ16, amphiphilic potential membrane forming peptide.
 XX
 KW Membrane-forming peptide; amphiphilic; monovalent metal cation; suture;
 KW internal lining; slow-diffusion drug delivery vehicle; erythropoietin;
 KW tissue-type plasminogen activator; haemoglobin; insulin; artificial skin;
 KW separation matrix; dialysis membranes; filter; collagen; cell migration;
 KW Alzheimer's disease; scrapie infection; therapy; proteolytic digestion;
 KW acidic pH resistant; stomach acid; cell adhesion; cell monolayer;
 KW tissue culture; differentiated cell; stratified cell layer.
 XX
 OS Synthetic.
 XX
 PN US5670483-A.
 PD 23-SEP-1997.
 XX
 PF 30-NOV-1994; 94US-00346849.
 XX
 PR 28-DEC-1992; 92US-00973326.
 XX
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX

XX Rich A, Holmes T, Zhang S, Lockshin C;
 XX WPI; 1997-479506/44.
 DR Membranes formed by self-assembly of amphiphilic peptide(s) - useful as
 PT bio:material(s), separation matrices, drug delivery vehicles, etc.
 XX
 PS Claim 30; Col 64; 49pp; English.
 CC AAW30219-W30252 are amphiphilic peptides used in the macroscopic membrane
 CC (MM) of the invention. The MM is formed by self-assembly of the
 CC amphiphilic peptides (with alternating hydrophilic and hydrophobic amino
 CC acids) in an aqueous medium containing monovalent metal cations. The MM
 CC are stable in serum, are non-cytotoxic, and are useful in biomaterial
 CC applications, such as medical products (e.g. sutures), or internal
 CC linings. The MM are useful as slow-diffusion drug delivery vehicles for
 CC protein-type drugs, including erythropoietin, tissue-type plasminogen
 CC activator, synthetic haemoglobin and insulin. They can be used in
 CC applications, such as separation matrices (e.g., dialysis membranes). The
 CC extremely small pore size (interfilament distance) of the MM makes them
 CC useful as filters. Collagen may be combined with the peptides to produce
 CC membranes for use as artificial skin. The MM may be used for making very
 CC thin, transparent fabric. Drugs which inhibit the self assembly of the
 CC peptides into filaments or filamentous membranes may be useful for
 CC treating Alzheimer's disease or scrapie infection. As they are resistant
 CC to proteolytic digestion and alkaline and acidic pH (such as stomach
 CC acid), drug delivery vehicles made of the MM could be taken orally. The
 CC charged residues and conformation of the MM promote cell adhesion and
 CC migration. The permeability of the MM also permits diffusion of small
 CC molecules, to the underside of cell monolayers, useful for tissue culture
 CC of differentiated cells and/or stratified cell layers. (Updated on 25-MAR
 CC -2003 to correct PF field.)
 XX
 SQ Sequence 16 AA;
 Query Match 78.6%; Score 66; DB 2; Length 16;
 Best Local Similarity 75.0%; Pred. No. 0.003;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 HNNNNNNNNNN 12
 Db 1 HNNHHNNHHNN 12
 RESULT 20
 AAY32989 standard; peptide; 16 AA.
 AC AAY32989;
 XX
 DT 09-NOV-1999 (first entry)
 DE Membrane forming aphiphilic peptide HON16.
 XX
 KM Membrane forming peptide; cell culturing; macroscopic membrane;
 KM amphiphilic peptide; slow-diffusion drug delivery system; cell growth;
 KM artificial skin; separation matrix; artificial tissue; scrapie infection;
 KM Alzheimer's disease; liver cirrhosis; kidney amyloidosis;
 KM protein conformational disease.
 XX
 OS Synthetic.
 XX
 PN US5955343-A.
 XX
 PD 21-SEP-1999.
 XX
 PF 22-AUG-1994; 94US-00293284.
 XX
 PR 28-DEC-1992; 92US-00973326.
 XX
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX

PI Rich A, Dipersio CM, Lockshin C, Holmes T, Zhang S;
 XX WPI; 1999-539576/45.
 DR Cell cultures utilizing stable macroscopic membranes formed by the self-
 PT assembly of amphiphilic peptides.
 XX
 PS Example 5; Col 27; 49pp; English.
 CC This sequence represents a membrane forming aphiphilic peptide used in
 CC the in vitro method of the invention. The method is for culturing cells
 CC utilizing stable macroscopic membranes formed by the self-assembly of
 CC amphiphilic peptides. The peptides are salt-induced to form insoluble and
 CC protease-resistant protein filaments with a beta-sheet secondary
 CC structure. The membranes may be useful in a wide variety of
 CC medical, research, industrial and biomaterial applications such as slow-
 CC diffusion drug delivery systems, artificial skin and separation matrices.
 CC The membranes may be used to support in vitro cell attachment and growth
 CC and for supporting artificial tissue (e.g. for in vivo use as implants).
 CC They are particularly useful as experimental models for Alzheimer's
 CC disease and scrapie infection and so may be used in disease modelling
 CC experiments and to assay for agents which modulate the disease processes.
 CC Additionally, they may be used in this way to study liver cirrhosis,
 CC kidney amyloidosis and other protein conformational diseases. The
 CC membrane in the membrane/cell mixture produced in the method: (i)
 CC supports cellular attachment and growth; (ii) self-assembles to form
 CC large, macroscopic membranes that are insoluble and stable in aqueous
 CC solutions, serum and ethanol; (iii) is highly resistant to heat,
 CC alkaline/acidic pHs, chemical denaturants and proteolytic digestion; (iv)
 CC is non-cytotoxic and non-immunogenic; (v) is visible to the naked eye if
 CC dyed but is otherwise transparent; (vi) may form thin, permeable, high
 CC density sheets or fibril like structures with simple structures, high
 CC tensile strength and a porous structure; (vii) may be metabolised by
 CC humans and animals; (viii) is inexpensive to produce; and (ix) can be
 CC produced and stored in sterile conditions
 XX
 SQ Sequence 16 AA;
 Query Match 78.6%; Score 66; DB 2; Length 16;
 Best Local Similarity 75.0%; Pred. No. 0.003;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 HNNNNNNNNNN 12
 Db 1 HNNHHNNHHNN 12
 RESULT 21
 AAY32988 standard; peptide; 16 AA.
 ID AAY32988
 AC AAY32988;
 XX
 DT 09-NOV-1999 (first entry)
 DE Membrane forming aphiphilic peptide HNO16.
 XX
 KM Membrane forming peptide; cell culturing; macroscopic membrane;
 KM amphiphilic peptide; slow-diffusion drug delivery system; cell growth;
 KM artificial skin; separation matrix; artificial tissue; scrapie infection;
 KM Alzheimer's disease; liver cirrhosis; kidney amyloidosis;
 KM protein conformational disease.
 XX
 OS Synthetic.
 XX
 PN US5955343-A.
 XX
 PD 21-SEP-1999.
 XX
 PF 22-AUG-1994; 94US-00293284.
 XX
 PR 28-DEC-1992; 92US-00973326.
 XX

PA (MASI) MASSACHUSETTS INST TECHNOLOGY.

XX Rich A, Dipersio CM, Lockshin C, Holmes T, Zhang S;

XX WPI, 1999-539576/45.

PT Cell cultures utilizing stable macroscopic membranes formed by the self-

XX assembly of amphiphilic peptides.

XX Example 5; Col 27; 49pp; English.

CC This sequence represents a membrane forming amphiphilic peptide used in
CC the in vitro method of the invention. The method is for culturing cells
CC utilizing stable macroscopic membranes formed by the self-assembly of
CC amphiphilic peptides. The peptides are salt-induced to form insoluble and
CC protease-resistant protein filaments with a beta-sheet secondary
CC structure. The membranes may be useful in a wide variety of
CC medical, research, industrial and biomaterial applications such as slow-
CC diffusion drug delivery systems, artificial skin and separation matrices.
CC The membranes may be used to support in vitro cell attachment and growth
CC and for supporting artificial tissue (e.g. for in vivo use as implants).
CC They are particularly useful as experimental models for Alzheimer's
CC disease and scrapie infection and so may be used in disease modelling
CC experiments and to assay for agents which modulate the disease processes.
CC Additionally, they may be used in this way to study liver cirrhosis,
CC kidney amyloidosis and other protein conformational diseases. The
CC membrane in the membrane/cell mixture produced in the method: (i)
CC supports cellular attachment and growth; (ii) self-assembles to form
CC large, macroscopic membranes that are insoluble and stable in aqueous
CC solutions, serum and ethanol; (iii) is highly resistant to heat,
CC alkaline/acidic pHs, chemical denaturants and proteolytic digestion; (iv)
CC is non-cytotoxic and non-immunogenic; (v) is visible to the naked eye if
CC density sheets or fibril like structures with simple structures, high
CC tensile strength and a porous structure; (vi) may be metabolised by
CC humans and animals; (vii) is inexpensive to produce; and (ix) can be
CC produced and stored in sterile conditions

XX SQ Sequence 16 AA;

Query Match 78.6%; Score 66; DB 2; Length 16;

Best Local Similarity 81.8%; Pred. No. 0.003;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 HNNNNNNNN 11

DB 1 HNNHNNHNNH 11

RESULT 22

AAE38886 standard; peptide; 16 AA.

AC AAE38886;

DT 18-DEC-2003 (first entry)

DE Membrane forming peptide, HON16.

KW Macroscopic membrane; biomaterial; medical product; internal lining;
KW artificial skin; slow-diffusion drug delivery; chromatography column;
KW separation matrix; suture; dialysis membrane; viral filter; fabric;
KW pathology; Alzheimer's disease; scrapie infection.

OS Unidentified.

PN US6548630-B1.

PD 15-APR-2003.

PF 22-JUL-1997; 97US-00898300.

PR 28-DEC-1992; 92US-00973326.

PR 30-NOV-1994; 94US-00346849.

XX (MASI) MASSACHUSETTS INST TECHNOLOGY.

XX Zhang S, Lockshin C, Rich A, Holmes T;

XX WPI; 2003-584339/55.

PT Macroscopic membrane useful in biomaterial applications e.g. sutures is
PT formed by self-assembly of amphiphilic peptides in an aqueous solution
XX containing monovalent metal cations.

XX Example 5; Col 25; 51pp; English.

CC The invention relates to a macroscopic membrane formed by self-assembly
CC of amphiphilic peptides in an aqueous solution containing monovalent
CC metal cations. The peptides have alternating hydrophobic and hydrophilic
CC amino acids and are complementary and structurally compatible. The
CC membranes are useful as biomaterial for medical products (e.g. sutures,
CC artificial skin, internal linings), as vehicles for slow-diffusion drug
CC delivery (preferably for protein type drugs e.g. erythropoietin, tissue
CC type plasminogen activator, synthetic haemoglobin and insulin), as
CC separation matrices (e.g. dialysis membranes, chromatography columns),
CC filters for the removal of viruses and other microscopic contaminants),
CC for other uses requiring permeable and water-insoluble material (e.g. for
CC culturing cell monolayers including differentiated cells and/or
CC stratified cell layers), for the preparation of very thin and transparent
CC fabric and as a model system for investigating the properties of
CC biological protein structures and providing insights into the pathology
CC of proteins and proteinaceous structures (e.g. Alzheimer's disease, scrapie
CC infection and in origin of life studies related to cell membranes and
CC cellular compartmentalisation). The present sequence is a membrane
XX forming peptide used in the invention

XX SQ Sequence 16 AA;

Query Match 78.6%; Score 66; DB 7; Length 16;

Best Local Similarity 75.0%; Pred. No. 0.003;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 HNNNNNNNNN 12

DB 1 HNNHNNHNNH 12

RESULT 23

AAE38885 standard; peptide; 16 AA.

AC AAE38885;

DT 18-DEC-2003 (first entry)

DE Membrane forming peptide, HNO16.

KW Macroscopic membrane; biomaterial; medical product; internal lining;
KW artificial skin; slow-diffusion drug delivery; chromatography column;
KW separation matrix; suture; dialysis membrane; viral filter; fabric;
KW pathology; Alzheimer's disease; scrapie infection.

OS Unidentified.

PN US6548630-B1.

PD 15-APR-2003.

PF 22-JUL-1997; 97US-00898300.

PR 28-DEC-1992; 92US-00973326.

PR 30-NOV-1994; 94US-00346849.

PA (MASI) MASSACHUSETTS INST TECHNOLOGY.

PI	Holmes T, Zhang S, Rich A, Dipersio CM, Lockshin C;
XX	
DR	WPI: 2004-707224/69.
XX	
PT	Culturing cells, in vitro, by adding macroscopic membrane formed by self
PT	assembly of amphiphilic peptides and monovalent metal cations, to form
PT	culture mixture, and maintaining culture mixture under conditions for
XX	cell growth.
PS	
XX	Example 5; SEQ ID NO 61; 50pp; English.
CC	
CC	The invention describes a method of culturing (M1) cells, in vitro. The
CC	cell culture involves adding a macroscopic membrane which is formed by
CC	self-assembly of amphiphilic peptides in an aqueous solution containing
CC	monovalent metal cations, where the peptides have alternating hydrophobic
CC	and hydrophilic amino acids and are complementary and structurally
CC	compatible, to a cell culture medium comprising cells, thus forming a
CC	membrane/culture mixture, and maintaining the mixture under conditions
CC	sufficient for cell growth. (M1) is useful for culturing a macroscopic
CC	membrane utilized in biomaterial applications e.g., slow-diffusion drug
CC	delivery systems, artificial skin or separation matrices, or as
CC	experimental models for Alzheimer's disease and scrapie infection e.g.,
CC	liver cirrhosis, kidney amyloidosis, or other protein conformational
CC	diseases. (M1) enables in vitro culturing of macroscopic membrane that is
CC	stable in aqueous solution, serum and ethanol, highly resistant to heat,
CC	alkaline and acidic pH, chemical denaturants and proteolytic digestion,
CC	and is non-cytotoxic. This is the amino acid sequence of an amphiphilic
CC	peptide that self-assembles into a membrane of the invention.
XX	
SQ	
	Sequence 16 AA;
Query Match	78.6%; Score 66; DB 8; Length 16;
Best Local Similarity	75.0%; Pred. NO. 0.003;
Matches	9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy	1 HHHNNNNNNNN 12
	1 HHHNNNNNNNN 12
Db	1 HOHHOHNNHOHN 12
RESULT 27	
ADSL5389	
ID	ADSL5389 standard; peptide; 16 AA.
XX	
AC	ADSL5389;
XX	
DT	16-DEC-2004 (first entry)
XX	
DE	
XX	Membrane forming amphiphilic peptide seqid 60.
KW	cell culture; macroscopic membrane; amphiphilic peptide; biomaterial;
KW	slow-diffusion drug delivery system; artificial skin; separation matrix;
KW	Alzheimer's disease; scrapie; liver cirrhosis; kidney amyloidosis;
KW	protein conformational disease; Yeast; DNA binding protein.
XX	
OS	Unidentified.
XX	
XX	
PN	US6800481-B1.
XX	
PD	05-OCT-2004.
XX	
PF	26-MAR-1997; 97US-00824513.
XX	
PR	28-DEC-1992; 92US-00973326.
XX	
PR	22-AUG-1994; 94US-00293284.
XX	
PA	(MASI) MASSACHUSETTS INST TECHNOLOGY.
PI	
XX	Holmes T, Zhang S, Rich A, Dipersio CM, Lockshin C;
XX	
DR	WPI: 2004-707224/69.

PT assembly of amphiphilic peptides and monovalent metal cations, to form
PT culture mixture, and maintaining culture mixture under conditions for
PT cell growth.
XX
PS Example 5; SEQ ID NO 60; 50bp; English.
XX
CC The invention describes a method of culturing (M1) cells, in vitro. The
CC cell culture involves adding a macroscopic membrane which is formed by
CC self-assembly of amphiphilic peptides in an aqueous solution containing
CC monovalent metal cations, where the peptides have alternating hydrophobic
CC and hydrophilic amino acids and are complementary and structurally
CC compatible, to a cell culture medium comprising cells, thus forming a
CC membrane/culture mixture, and maintaining the mixture under conditions
CC sufficient for cell growth. (M1) is useful for culturing a macroscopic
CC membrane utilised in biomaterial applications e.g., slow-diffusion drug
CC delivery systems, artificial skin or separation matrices, or as
CC experimental models for Alzheimer's disease and scrapie infection e.g.,
CC liver cirrhosis, kidney amyloidosis, or other protein conformational
CC diseases. (M1) enables in vitro culturing of macroscopic membrane that is
CC stable in aqueous solution, serum and ethanol, highly resistant to heat,
CC alkaline and acidic pH, chemical denaturants and proteolytic digestion,
CC and is non-cytotoxic. This is the amino acid sequence of an amphiphilic
CC peptide that self-assembles into a membrane of the invention.
XX
SQ Sequence 16 AA;
XX
Query Match 78.6%; Score 66; DB 8; Length 16;
Best Local Similarity 81.8%; Pred. No. 0.003; Mismatches 0; Gaps 0;
Matches 9; Conservative 0; Indels 2; Indels 0; Gaps 0;
XX
QY 1 HNNNNNNNNH 11
1 HNNNNNNNNH 11
XX
Db 1 HNNNNNNNNH 11
XX
RESULT 28
AAB59190
ID AAB59190 standard; protein; 124 AA.
XX
AC AAB59190;
XX
DT 22-MAR-2001 (first entry)
XX
DE Mouse NADE.
XX
KW Neurotrophin receptor, p75-NTR; NGF-induced apoptosis;
KW neurogenetic disease; NF-kappaB.
XX
OS Mus sp.
XX
PN MO200075278-A2.
XX
PD 14-DEC-2000.
XX
PF 07-JUN-2000; 2000MO-US015621.
XX
PR 07-JUN-1999; 99US-00327750.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Sato T;
XX
DR WPI; 2001-061707/07.
XX
PT New p75-neurotrophin receptor-associated cell death executor (NADE) and
PT the gene encoding NADG, useful for modulating the activity of p75NTR and
PT for detecting neurodegenerative diseases.
XX
PS Disclosure; Fig 1; 134pp; English.
XX
CC The present invention relates to a purified polypeptide capable of
CC binding neurotrophin receptor (p75-NTR). The invention is useful for
CC binding and modulating the activity of p75NTR. The peptide mediates NGF-

CC induced apoptosis, which plays an important role in neurogenetic
CC diseases. The peptide of the invention and p75NTR are useful for
CC inhibiting NF-kappaB activation in a cell or a subject, for inducing
CC caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase
CC and fragment nuclear DNA in a cell by co-expression of (I) and p75-NTR
XX
SQ Sequence 124 AA;
XX
Query Match 77.4%; Score 65; DB 4; Length 124;
Best Local Similarity 75.0%; Pred. No. 0.037; Mismatches 3; Indels 0; Gaps 0;
Matches 9; Conservative 3; Indels 0; Gaps 0;
XX
QY 1 HNNNNNNNNH 12
1 HNNNNNNNNH 12
XX
Db 37 NNNNNNNNNH 48
XX
RESULT 29
ABB48454
ID ABB48454 standard; protein; 303 AA.
XX
AC ABB48454;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #1158.
XX
KW Antibacterial, gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; Bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN MO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001MO-FR001118.
XX
PR 11-APR-2000; 2000FR-00004629.
XX
PA (INSP) INST PASTEUR.
XX
PI Buchrieser C, Frangeul L, Couve E, Ruenick C, Faini H, Dehoux P,
PI Bussarguet O, Cherouani F, Nedjari H, Glaeser P, Kunst F, Cossart P,
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA,
PI Dominguez-Bernal G, Garrido-Garcia P, Hierrez-Martinez A, Amend A,
PI Chakraborty T, Doman E, Hain T, Berche P, Chabdit A, Durant L,
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N,
PI Madueno E, De Pablo B, Weiland J, Kaerst U, Entian K, Hauf J,
PI Rose M, Voess H;
XX
DR WPI; 2002-010914/01.
XX
PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and related
PT polypeptides.
XX
PS Claim 6; SEQ ID NO 1159; 192pp; French.
XX
CC The present invention relates to the genome sequence of Listeria
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Protein
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of L. monocytogenes and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by L.

CC monocytogenes and related organisms. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

Sequence 303 AA;

Query Match 77.4%; Score 65; DB 5; Length 303;
Best Local Similarity 75.0%; Pred. No. 0.097;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HHHNNHHNNH 12
Db 3 HHHDHAGHHNN 14

RESULT 30
AAG47764

ID AAG47764 standard; protein: 167 AA.

AC AAG47764;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 60237.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EF1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 29-MAR-1999; 99US-0126264P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 21-APR-1999; 99US-0130077P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 04-MAY-1999; 99US-0132048P.

PR 05-MAY-1999; 99US-0132484P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 14-JUN-1999; 99US-0138847P.

PR 16-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 17-JUN-1999; 99US-0139453P.

PR 18-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 21-JUN-1999; 99US-0139763P.

PR 22-JUN-1999; 99US-0139817P.

PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140658P.

PR 28-JUN-1999; 99US-0140823P.

PR 29-JUN-1999; 99US-0140991P.

PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.

PR 02-JUL-1999; 99US-0142154P.

PR 06-JUL-1999; 99US-0142055P.

PR 08-JUL-1999; 99US-0142390P.

PR 09-JUL-1999; 99US-0142803P.

PR 12-JUL-1999; 99US-0142920P.

PR 13-JUL-1999; 99US-0142977P.

PR 14-JUL-1999; 99US-0143542P.

PR 15-JUL-1999; 99US-0143624P.

PR 16-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.

PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.

PR 20-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144633P.

PR 21-JUL-1999; 99US-0144884P.

PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.

PR 22-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145089P.

PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.

PR 23-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.

PR 26-JUL-1999; 99US-0145224P.

PR 27-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.

PR 28-JUL-1999; 99US-0145919P.

PR 02-AUG-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.

PR 03-AUG-1999; 99US-0146389P.

PR 04-AUG-1999; 99US-0147038P.

PR 04-AUG-1999; 99US-0147204P.

PR 04-AUG-1999; 99US-0147302P.

PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 76.2%; Score 64; DB 3; Length 167;
Best Local Similarity 75.0%; Pred. No. 0.07;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 HNNNNNNNNNN 12
: : : : :
Db 96 NNNSHNNNNNN 107

RESULT 31
AAG20600
ID AAG20600 standard; protein, 193 AA.
XX
AC AAG20600;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22855.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridization assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0132484P.
PR 23-APR-1999; 99US-0132485P.
PR 28-APR-1999; 99US-0133149P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135533P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.

```
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 21-JUN-1999; 99US-0139763P.
PR 22-JUN-1999; 99US-0139817P.
PR 23-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140824P.
PR 30-JUN-1999; 99US-0140991P.
PR 01-JUL-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145918P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149375P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150565P.
PR 26-AUG-1999; 99US-0150565P.
PR 27-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151438P.
PR 07-SEP-1999; 99US-0151930P.
PR 10-SEP-1999; 99US-0152363P.
PR 13-SEP-1999; 99US-0153070P.
PR 15-SEP-1999; 99US-0153758P.
PR 16-SEP-1999; 99US-0154018P.
PR 20-SEP-1999; 99US-0154039P.
PR 22-SEP-1999; 99US-0154779P.
PR 23-SEP-1999; 99US-0155139P.
PR 24-SEP-1999; 99US-0155486P.
PR 28-SEP-1999; 99US-0156599P.
PR 29-SEP-1999; 99US-0156488P.
PR 04-OCT-1999; 99US-0156596P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157753P.
PR 07-OCT-1999; 99US-0157865P.
PR 08-OCT-1999; 99US-0158029P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159310P.
PR 14-OCT-1999; 99US-0159311P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159684P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161982P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149375P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150565P.
PR 26-AUG-1999; 99US-0150565P.
PR 27-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151438P.
PR 07-SEP-1999; 99US-0151930P.
PR 10-SEP-1999; 99US-0152363P.
PR 13-SEP-1999; 99US-0153070P.
PR 15-SEP-1999; 99US-0153758P.
PR 16-SEP-1999; 99US-0154018P.
PR 20-SEP-1999; 99US-0154039P.
PR 22-SEP-1999; 99US-0154779P.
PR 23-SEP-1999; 99US-0155139P.
PR 24-SEP-1999; 99US-0155486P.
PR 28-SEP-1999; 99US-0156599P.
PR 29-SEP-1999; 99US-0156488P.
PR 04-OCT-1999; 99US-0156596P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157753P.
PR 07-OCT-1999; 99US-0157865P.
PR 08-OCT-1999; 99US-0158029P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159310P.
PR 14-OCT-1999; 99US-0159311P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159684P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161982P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match

Best Local Similarity 76.2%; Score 64; DB 3; Length 193;

Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Cv 1 HNNHHNNHHN 12

Db 122 NHHSHNNHHN 133

RESULT 32
AAG20599
ID AAG20599 standard; protein; 244 AA.
AC AAG20599;
XX
XX
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 22854.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135533P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138647P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142877P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144252P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.

PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149732P.
PR	23-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149930P.
PR	23-AUG-1999;	99US-0149932P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	30-AUG-1999;	99US-0151080P.
PR	31-AUG-1999;	99US-0151303P.
PR	01-SEP-1999;	99US-0151438P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157153P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158365P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	18-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159841P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160981P.
PR	23-OCT-1999;	99US-0160981P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	26-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	28-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	29-OCT-1999;	99US-0161933P.
PR	29-OCT-1999;	99US-0162142P.

```

Query Match          76.2%; Score 64; DB 3; Length 244;
Best Local Similarity 75.0%; Pred. No. 0.11;
Matches          9; Conservative          2; Mismatches          1; Indels          0; Gaps          0.

QY          1 HHHHHHHHHHHN 12
              :|||:|||||
Db          173 NNSHSHQNQNNHN 184

RESULT 33
ID          AAG47763 standard; protein; 244 AA.
XX          AAG47763;

```

XX	DT	18-OCT-2000	(first entry)
XX	XX	Arabidopsis thaliana protein fragment SEQ ID NO: 60236.	
DE	XX	Protein identification; signal transduction pathway; metabolic pathway;	
KW	XX	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	XX	termination sequence.	
OS	XX	Arabidopsis thaliana.	
PN	XX	EP1033405-A2.	
XX	PD	06-SEP-2000.	
XX	PF	25-FEB-2000; 2000EP-00301439.	
XX	PR	25-FEB-1999; 99US-0121825P.	
PR	PR	05-MAR-1999; 99US-0123180P.	
PR	PR	09-MAR-1999; 99US-0123548P.	
PR	PR	23-MAR-1999; 99US-0125788P.	
PR	PR	25-MAR-1999; 99US-0126264P.	
PR	PR	29-MAR-1999; 99US-0126765P.	
PR	PR	01-APR-1999; 99US-0127463P.	
PR	PR	06-APR-1999; 99US-0128234P.	
PR	PR	08-APR-1999; 99US-0128714P.	
PR	PR	16-APR-1999; 99US-0129845P.	
PR	PR	19-APR-1999; 99US-0130077P.	
PR	PR	21-APR-1999; 99US-0130449P.	
PR	PR	23-APR-1999; 99US-0130510P.	
PR	PR	28-APR-1999; 99US-0130891P.	
PR	PR	30-APR-1999; 99US-0131449P.	
PR	PR	30-APR-1999; 99US-0132048P.	
PR	PR	04-MAY-1999; 99US-0132407P.	
PR	PR	05-MAY-1999; 99US-0132484P.	
PR	PR	06-MAY-1999; 99US-0132485P.	
PR	PR	06-MAY-1999; 99US-0132486P.	
PR	PR	07-MAY-1999; 99US-0132487P.	
PR	PR	11-MAY-1999; 99US-0132863P.	
PR	PR	14-MAY-1999; 99US-0134256P.	
PR	PR	14-MAY-1999; 99US-0134218P.	
PR	PR	14-MAY-1999; 99US-0134219P.	
PR	PR	14-MAY-1999; 99US-0134221P.	
PR	PR	18-MAY-1999; 99US-0134370P.	
PR	PR	19-MAY-1999; 99US-0134768P.	
PR	PR	20-MAY-1999; 99US-0134941P.	
PR	PR	21-MAY-1999; 99US-0135533P.	
PR	PR	24-MAY-1999; 99US-0135629P.	
PR	PR	25-MAY-1999; 99US-0136021P.	
PR	PR	27-MAY-1999; 99US-0136392P.	
PR	PR	28-MAY-1999; 99US-0136782P.	
PR	PR	01-JUN-1999; 99US-0137222P.	
PR	PR	03-JUN-1999; 99US-0137528P.	
PR	PR	04-JUN-1999; 99US-0137502P.	
PR	PR	07-JUN-1999; 99US-0137724P.	
PR	PR	08-JUN-1999; 99US-0138094P.	
PR	PR	10-JUN-1999; 99US-0138540P.	
PR	PR	10-JUN-1999; 99US-0138847P.	
PR	PR	14-JUN-1999; 99US-0139119P.	
PR	PR	16-JUN-1999; 99US-0139452P.	
PR	PR	16-JUN-1999; 99US-0139453P.	
PR	PR	17-JUN-1999; 99US-0139454P.	
PR	PR	18-JUN-1999; 99US-0139455P.	
PR	PR	18-JUN-1999; 99US-0139456P.	
PR	PR	18-JUN-1999; 99US-0139457P.	
PR	PR	18-JUN-1999; 99US-0139458P.	
PR	PR	18-JUN-1999; 99US-0139459P.	
PR	PR	18-JUN-1999; 99US-0139460P.	
PR	PR	18-JUN-1999; 99US-0139461P.	
PR	PR	18-JUN-1999; 99US-0139462P.	
PR	PR	18-JUN-1999; 99US-0139463P.	

PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 25-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.

PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151388P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153708P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161044P.
PR 25-OCT-1999; 99US-0161045P.
PR 25-OCT-1999; 99US-0161065P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 76.2%; Score 64; DB 3; length 244;
Best Local Similarity 75.0%; Pred. No. 0.11;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
Db 173 NNHSHNNHNN 184

RESULT 34
AD061793
ID AD061793 standard; protein; 244 AA.

XX ADO61793;
XX
XX 15-JUL-2004 (first entry)

DE Transcription factor G1591, SEQ ID 260.

XX Plant; transcription factor; transgenic plant; abiotic stress tolerance;
XX osmotic stress tolerance; cold tolerance; heat tolerance;
KW

KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
 KW glyophosate resistance; flowering; fertility; seed development.
 OS Arabidopsis thaliana.
 XX
 PN WO2004031349-A2.
 XX
 XX 15-APR-2004.
 XX
 PD 18-SEP-2003; 2003WO-US030292.
 XX
 PR 18-SEP-2002; 2002US-0411837P.
 PR 17-DEC-2002; 2002US-043416P.
 PR 24-APR-2003; 2003US-0465809P.
 XX
 PA (MENDEL-) MENDEL BIOTECHNOLOGY INC.
 XX
 PI Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL,
 PI Reichmann JT, Haake V, Dubell AN, Keddie JS, Sherman BK,
 XX WPI; 2004-330163/30.
 DR N-PSDB; AD061792.
 XX
 PT New recombinant polynucleotide encoding transcription factor
 PT polypeptides, useful for producing transgenic plants with advantageous
 XX properties compared to a reference plant.
 PS Disclosure; SEQ ID NO 260; 510pp; English.
 XX
 CC The present invention relates to novel plant transcription factor
 CC proteins (i) and nucleotide sequences (ii) (AD061534-AD063778). The
 CC sequences can be used to produce transgenic plants, which overexpress
 CC (ii), where the transgenic plant has an altered trait as compared to a
 CC non-transgenic plant or wild-type plant. The transgenic plant comprises
 CC an altered trait selected from increased tolerance to abiotic stresses,
 CC increased tolerance to osmotic stress, increased tolerance to cold,
 CC increased germination in cold, increased tolerance to heat, increased
 CC germination in heat, increased tolerance to freezing conditions,
 CC increased tolerance to low nitrogen conditions, increased tolerance to
 CC low phosphate conditions, increased tolerance to disease, including
 CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
 CC glyophosate, increased sensitivity to ABA, reduced sensitivity to ABA,
 CC to sugars, altered carbon/nitrogen sensing, early flowering, increased tolerance
 CC to altered shoot meristem development, altered branching pattern, reduced
 CC fertility, altered flower structure, loss of flower determinacy, late
 CC altered stem morphology, altered vascular tissue structure, reduced
 CC apical dominance, altered trichome density, altered trichome development,
 CC altered trichome structure, altered root development, altered shade
 CC avoidance, altered seed development, altered seed ripening, altered seed
 CC germination, slow growth, fast growth, altered cell differentiation,
 CC altered cell proliferation, altered cell expansion, altered phase change,
 CC altered senescence, abnormal embryo development, altered programmed cell
 CC death, lethality when overexpressed, altered necrosis patterns, increased
 CC plant size, increased biomass, large seedlings, dwarfed plants, increased
 CC green leaves, change in leaf shape, increased leaf size and mass, light
 CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
 CC seed, increased seed coloration, altered seed size, altered seed shape, large
 CC content, altered seed protein content, altered seed fatty acids, altered seed oil
 CC altered leaf prenyl lipid content, increased seedprenyl content,
 CC decreased anthocyanin levels. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 244 AA;

Query Match 76.2%; Score 64; DB 8; Length 244;
 Best Local Similarity 66.7%; Pred. No. 0.11;
 Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 HHHHHHHHHHHH 12
 DB 115 HHHHHHHHHHHH 126
 RESULT 35
 AAG47762
 ID AAG47762 standard; protein; 300 AA.
 XX
 AC AAG47762;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 60235.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridization assay; genetic mapping; gene expression control; promoter;
 XX termination sequence.
 OS Arabidopsis thaliana.
 XX
 PN EP103405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-00301439.
 XX
 PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126254P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 28-APR-1999; 99US-0130891P.
 PR 30-APR-1999; 99US-0131449P.
 PR 04-MAY-1999; 99US-0132407P.
 PR 05-MAY-1999; 99US-0132484P.
 PR 06-MAY-1999; 99US-0132485P.
 PR 07-MAY-1999; 99US-0132486P.
 PR 11-MAY-1999; 99US-0132637P.
 PR 14-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 18-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 20-MAY-1999; 99US-0134941P.
 PR 21-MAY-1999; 99US-0135124P.
 PR 24-MAY-1999; 99US-0135353P.
 PR 25-MAY-1999; 99US-0136029P.
 PR 28-MAY-1999; 99US-0136021P.
 PR 28-MAY-1999; 99US-0136392P.
 PR 01-JUN-1999; 99US-0136782P.
 PR 03-JUN-1999; 99US-0137222P.
 PR 04-JUN-1999; 99US-0137528P.
 PR 07-JUN-1999; 99US-0137502P.
 PR 08-JUN-1999; 99US-0137724P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 17-JUN-1999; 99US-0139453P.
 PR 99US-0139492P.

ID AAG20598 standard; protein; 301 AA.
AC AAG20598;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22853.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-01233180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132488P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134376P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135233P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136382P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137222P.
PR 04-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139119P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140655P.
PR 26-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 09-JUL-1999; 99US-0142803P.
PR 12-JUL-1999; 99US-0142920P.
PR 13-JUL-1999; 99US-0142977P.
PR 14-JUL-1999; 99US-0143542P.
PR 15-JUL-1999; 99US-0144065P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145918P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146386P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147453P.
PR 10-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.


```
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150864P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 76.2%; Score 64; DB 3; Length 301;
Best Local Similarity 75.0%; Pred. No. 0.13; Indels 0; Gaps 0;
Matches 9; Conservative 2; Mismatches 1;

QY 1 HNNNNNNNNNN 12
DB 230 NNNSHNNNNNN 241

RESULT 37
ABU19286
ID ABU19286 standard; protein; 451 AA.
XX
AC ABU19286;
XX
XX 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #4813.

```
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
KM Borrelia burgdorferi.
OS WO200277183-A2.
XX 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELITR) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACA23156.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 47210; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation; (7) identifying a gene in an operon required for
CC proliferation; (8) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 451 AA;
```

Query Match 76.2%; Score 64; DB 6; Length 451;
Best Local Similarity 66.7%; Pred. No. 0.21;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 HNNNNNNNNNN 12
DB 253 HDHHDHNDHNDH 264

RESULT 38

AAG47714

ID AAG47714 standard; protein; 913 AA.

AC AAG47714;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 60168.

XX Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 04-MAY-1999; 99US-0132407P.

PR 05-MAY-1999; 99US-0132484P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132487P.

PR 11-MAY-1999; 99US-0132863P.

PR 14-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 18-MAY-1999; 99US-0134370P.

PR 19-MAY-1999; 99US-0134370P.

PR 20-MAY-1999; 99US-0134768P.

PR 21-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137222P.

PR 04-JUN-1999; 99US-0137528P.

PR 07-JUN-1999; 99US-0137528P.

PR 08-JUN-1999; 99US-0137724P.

PR 10-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 14-JUN-1999; 99US-0138847P.

PR 16-JUN-1999; 99US-0139119P.

PR 17-JUN-1999; 99US-0139452P.

PR 17-JUN-1999; 99US-0139453P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139751P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139817P.
PR 23-JUN-1999; 99US-0140053P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 09-JUL-1999; 99US-0142803P.
PR 12-JUL-1999; 99US-0142927P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145242P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146386P.
PR 03-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147416P.
PR 10-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148311P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.

PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-015717P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 18-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160880P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 76.2%; Score 64; DB 3; Length 913;
Best Local Similarity 75.0%; Pzed. No. 0.44;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHHNNHHN 12
Db 394 NNHSHNNHHN 405

RESULT 39
AAG47713
ID AAG47713 standard; protein; 923 AA.
XX
AC AAG47713;
XX

DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 60167.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
PN EPI033405-A2.
PD
XX 06-SEP-2000.
PF 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134707P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134841P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.

```
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 28-JUN-1999; 99US-0140695P.  
PR 29-JUN-1999; 99US-0140823P.  
PR 30-JUN-1999; 99US-0140991P.  
PR 01-JUL-1999; 99US-0141287P.  
PR 02-JUL-1999; 99US-0141842P.  
PR 06-JUL-1999; 99US-0142154P.  
PR 08-JUL-1999; 99US-0142055P.  
PR 09-JUL-1999; 99US-0142380P.  
PR 12-JUL-1999; 99US-0142803P.  
PR 13-JUL-1999; 99US-0142920P.  
PR 14-JUL-1999; 99US-0142977P.  
PR 15-JUL-1999; 99US-0143542P.  
PR 16-JUL-1999; 99US-0143624P.  
PR 16-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 19-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 21-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145221P.  
PR 26-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 04-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 08-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 10-AUG-1999; 99US-0147935P.  
PR 11-AUG-1999; 99US-0148171P.  
PR 12-AUG-1999; 99US-0148319P.  
PR 13-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149368P.  
PR 18-AUG-1999; 99US-0149175P.  
PR 20-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 23-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0149930P.  
PR 26-AUG-1999; 99US-0150566P.  
PR 27-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0152367P.  
PR 13-SEP-1999; 99US-0153700P.  
PR 15-SEP-1999; 99US-0153758P.  
PR 16-SEP-1999; 99US-0154018P.  
PR 20-SEP-1999; 99US-0154039P.  
PR 22-SEP-1999; 99US-0154779P.  
PR 23-SEP-1999; 99US-0155139P.  
PR 24-SEP-1999; 99US-0155486P.  
PR 28-SEP-1999; 99US-0155659P.  
PR 29-SEP-1999; 99US-0156558P.  
PR 04-OCT-1999; 99US-0156566P.  
PR 05-OCT-1999; 99US-0157117P.  
PR 06-OCT-1999; 99US-0157153P.  
PR 07-OCT-1999; 99US-0157865P.  
PR 08-OCT-1999; 99US-0158029P.  
PR 12-OCT-1999; 99US-0158233P.  
PR 13-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 14-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 21-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 25-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161040P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 26-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.  
  
Query Match 76.2%; Score 64; DB 3; Length 923;  
Best Local Similarity 75.0%; Pred. No. 0.44;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 HHNNNNNNNN 12  
DB 404 NNHSHNNNNNN 415  
  
RESULT 40  
AA647712  
ID AA647712 standard; protein; 993 AA.  
XX  
AC AA647712;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 60166.  
DE Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.
```

XX Arabidopsis thaliana.
OS
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.

PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145318P.
PR 27-JUL-1999; 99US-0145319P.
PR 28-JUL-1999; 99US-0145351P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.

Accession	Protein Name	Organism	Function
RESULT 41			
ADM98847			
ID	ADM98847	standard	protein, 1176 AA.
XX			
AC	ADM98847		
XX			
DT	01-JUL-2004	(first entry)	
XX			
DE	HMG-CoA reductase polypeptide #100.		
XX			
KW	Geranylgeranyl pyrophosphate synthase; diterpene; diterpene precursor		
KW	diterpene synthase; defence toxin; volatile defensive signal;		
KW	pollinator attractant; photoprotectant; HMG-CoA reductase; enzyme.		
XX			
OS	Phycomyces blakesleeanus.		
XX			
EN	US2004072323-A1.		
XX			
PD	15-APR-2004.		
XX			

The invention relates to a unicellular organism for producing a diterpene or diterpene precursor comprising an exogenous nucleic acid sequence encoding a geranylgeranyl pyrophosphate synthase under the control of a promoter operable in the organism, and an exogenous nucleic acid sequence encoding a diterpene synthase under the control of a promoter operable in the organism. The invention also relates to methods of producing a diterpene or diterpene precursor and a method of isolating a diterpene synthase comprising growing several cells in the presence of a polyomeric resin to make a cell/resin mixture, where at least one of the cells further comprises at least one isolated and purified nucleic acid sequence of a yeast expression library, and the expression of the nucleic acid sequence is regulated by an inducible promoter under conditions where the expression is induced, filtering the cell/resin mixture, extracting the cell/resin mixture with alcohol to produce an organic eluent and analysing the organic eluent by a screening method including chromatography and/or spectroscopy, to identify the nucleic acid sequence encoding the diterpene synthase. The unicellular microorganism is useful as a diterpene or diterpene precursor producing system. Diterpenes, in plants, serve as defence toxins, volatile defensive signals, pollinator attractants and photoprotectants. This sequence represents an HMG-CoA reductase polypeptide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

RESULT	42
ID	AAM58573
XX	AAM58573 standard; protein; 341 AA.
XX	
AC	AAM58573;
XX	
DT	17-OCT-2003 (revised)
DT	25-MAR-2003 (revised)
DT	12-OCT-1998 (first entry)
XX	
DE	Trichoderma reesei ACEII transcriptional activator protein.
XX	
KW	ACEII; transcriptional activator protein; ace2 gene; cellulase;
KW	cellobiohydrolase; endoglucanase; beta-glucosidase; cellulase;
KW	alpha-arabinosidase; alpha-D-glucuronidase; acetyl esterases; mannanase
KW	pectinase; pectinesterase; pectin acid lyase.
OS	Hypocrea jecorina; strain Rut-C-30.
XX	
Key	Location/Organisms
PH	

```

FT      Domain
FT      /note="zn2c6 binuclear cluster domain"
FT      Region
FT      47..65
FT      /note="His-rich region homologous to Drosophila homeobox
FT      -containing regulatory protein"
FT      Region
FT      81..99
FT      /note="Gln/Pro-rich region homologous to Drosophila
FT      regulatory protein"
XX      MO9823642-A1.
XX      04-JUN-1998.
XX      01-DEC-1997; 97MO-FI000743.
XX      29-NOV-1996; 96US-0032156P.
XX      13-DEC-1996; 96US-0032959P.
XX      10-MAR-1997; 97US-0040140P.
XX      (ROHG ) ROEHM ENZYME FINLAND OY.
XX      Saloheimo A, Aro N, Ilmen M, Penttilae M;
XX      WPI, 1998-322661/28.
XX      N-PSDB; AAV31355, AAV31356.
XX      New isolated transcriptional regulatory protein nucleic acid(s) - used
XX      to, e.g. enhance production of homologous and heterologous proteins in
XX      fungal hosts.
XX      Claim 29; Page 60; 99pp; English.
XX      This polypeptide comprises the novel ACEII transcriptional activator
XX      protein of Trichoderma reesei Rut-C-30 that is capable of activating the
XX      promoter of the cchl gene that encodes a major cellulase
XX      cellobiohydrolase I protein. It has features indicative of a regulatory
XX      protein (see AAW58575-77). cDNA and genomic DNA sequences (see AAV31355
XX      and AAV31356) that code for ACEII are provided, as well as related ACEI
XX      (see AAW58572). The invention provides a method of stimulating gene
XX      expression in hosts transformed with sequences encoding aceI and ace2 by
XX      providing to the host a DNA construct in which the gene of interest is
XX      operably linked to a promoter that further contains one or more binding
XX      sites for the aceI and/or ace2 transcriptional activator proteins that
XX      are heterologous to the native promoter structure. Also provided is a
XX      method for enhancing expression of a desired gene in cells capable of
XX      expressing ACEI and/or ACEII by inserting into the promoter of the gene a
XX      binding site for ACEI and/or ACEII, or multiple copies of such sites. The
XX      methods can be used to enhance expression in fungal cells of homologous
XX      or heterologous proteins such as enzymes, e.g. cellobiohydrolase I or II,
XX      endoglucanase I, II, III and V, beta-glucosidase, xylanase, alpha-
XX      arabinosidase, alpha-D-glucuronidase, acetyl esterase, mannanase,
XX      pectinase, pectinesterase, and pectin acid lyase (all claimed). (Updated
XX      on 25-MAR-2003 to correct PA field.) (Updated on 17-OCT-2003 to
XX      standardise OS field)
XX      SQ      Sequence 341 AA;
XX      Query Match      75.0%; Score 63; DB 2; Length 341;
XX      Best Local Similarity 66.7%; Pred. No. 0.21;
XX      Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
XX      QY      1 HNNHNNHNNHNN 12
XX      |:|:|:|:|:|
XX      DB      55 HESHSHSHNNH 66
XX      RESULT 43
XX      ID      AAY36191 standard; protein; 150 AA.
XX      AC      AAY36191;
XX      DT      23-SEP-1999 (first entry)

```

```

XX      Human secreted protein #63.
XX      DE      Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;
XX      KW      diagnostic; gene therapy; chromosome mapping; secretion vector.
XX      OS      Homo sapiens.
XX      PN      MO925825-A2.
XX      27-MAY-1999.
XX      13-NOV-1998; 98WO-IB001862.
XX      13-NOV-1997; 97US-0066677P.
XX      17-DEC-1997; 97US-0069957P.
XX      09-FEB-1998; 98US-0074121P.
XX      13-APR-1998; 98US-0081563P.
XX      10-SEP-1998; 98US-0096116P.
XX      04-SEP-1998; 98US-0099273P.
XX      (GEST ) GENSET.
XX      PI      Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX      WPI; 1999-347472/29.
XX      DR      N-PSDB; AAX97875.
XX      PT      Extended cDNAs encoding secreted proteins.
XX      PS      Claim 7; Page 293; 307pp; English.
XX      AAY36129-Y36222 represent novel human secreted proteins encoded by the
XX      extended cDNA sequences represented in AAX97813-X97906. The proteins of
XX      the invention have cytosolic, thrombotic and osteopathic activity. The
XX      extended cDNAs can be used to express secreted proteins or parts of them
XX      or to obtain antibodies capable of binding to the secreted proteins. They
XX      may also be used in diagnostic, forensic, gene therapy and chromosome
XX      mapping procedures. Uses also include design of expression vectors and
XX      secretion vectors
XX      SQ      Sequence 150 AA;
XX      Query Match      73.8%; Score 62; DB 2; Length 150;
XX      Best Local Similarity 72.7%; Pred. No. 0.12;
XX      Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX      QY      1 HNNHNNHNNH 11
XX      |:|:|:|:|
XX      DB      124 HSHNHQSHNNH 134
XX      RESULT 44
XX      ID      ADJ46043 standard; protein; 150 AA.
XX      AC      ADJ46043;
XX      DT      06-MAY-2004 (first entry)
XX      DE      Novel human secreted protein-related protein sequence Segid196.
XX      KW      secreted protein; upstream regulator; gene therapy; protein purification;
XX      protein synthesis; chromosomal mapping; individual identification;
XX      forensic; hereditary disease; drug reaction; immunosassay;
XX      epitope mapping; vaccine; immune system regulation;
XX      haematopoietic system; tissue growth; reproductive hormone;
XX      cell migration; blood clotting; receptor/ligand interaction;
XX      adhesion molecule; assisted drug delivery;
XX      human glial maturation factor gamma-2; neurite outgrowth;
XX      neurite resprouting; human.
XX      OS      Homo sapiens.

```

XX US2003144490-A1.
 PN 31-JUL-2003.
 XX 10-DEC-2002; 2002US-00319763.
 XX 13-NOV-1997; 97US-0066677P.
 PR 17-DEC-1997; 97US-0069957P.
 PR 09-FEB-1998; 98US-0074121P.
 PR 13-APR-1998; 98US-0081563P.
 PR 10-AUG-1998; 98US-0096116P.
 PR 04-SEP-1998; 98US-0099273P.
 PR 13-NOV-1998; 98US-00191997.
 PR 15-SEP-2000; 2000US-00663600.
 XX (EDWA/) EDWARDS J D M.
 PA (DUC/) DUCLET A.
 XX (BOUG/) BOUGUELERET L.
 PI Edwards JDM, Duclet A, Bougueleret L;
 XX WPI; 2003-851788/79.
 DR N-PSDB; ADU45996.
 XX New nucleic acid encoding secreted human polypeptides, useful e.g. in
 PT gene therapy or diagnosis, also encoded proteins, potential therapeutic
 PT agents.
 PS Claim 1; SEQ ID NO 196; 269pp; English.
 XX This invention relates to novel purified isolated polynucleotides which
 CC comprise a sequence that encodes at least 10 amino acids (aa) from any of
 CC 48 secreted polypeptide sequences, given in the specification, or
 CC fragments of polypeptides encoded by human cDNA contained in the
 CC corresponding deposited clone. The DNA sequences of the invention encode
 CC secreted proteins (or their fragments), and can be used to
 CC identify/isolate upstream regulators, potentially useful in gene therapy
 CC or protein purification, by controlling protein synthesis, as probes for
 CC chromosomal mapping, identification of individuals, and for diagnosis or
 CC forensics, for example identifying genes associated with hereditary
 CC diseases or drug reactions, for recombinant expression of the encoded
 CC proteins or, where the DNA sequence encodes a signal peptide, for
 CC directing secretion of heterologous polypeptides. Polypeptides encoded by
 CC the DNA sequences of the invention can be used to raise antibodies,
 CC anti-idiotypic antibodies, as tags in for example immunoassays, epitope
 CC mapping or vaccines, also as molecular weight markers, to screen for
 CC agents with biological activity and as therapeutic agents with,
 CC potentially, a very wide range of activities, for example regulation of
 CC the immune or haematopoietic systems, tissue growth, reproductive
 CC hormones, cell migration, blood clotting or receptor/ligand interaction
 CC also as adhesion molecules for assisted drug delivery. A typical isolated
 CC sequence is human glial maturation factor gamma-2, which stimulates
 CC neurite outgrowth and resprouting. The present sequence is that of a
 CC human secreted protein of the invention.
 XX Sequence 150 AA;
 SQ
 Query Match 73.8%; Score 62; DB 7; Length 150;
 Best Local Similarity 72.7%; Pred. No. 0.12;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HNNHNNHNNH 11
 Db 124 HSHNHSHSHNH 134
 RESULT 45
 ADP19452
 ID ADP19452 standard; protein, 150 AA.
 XX ADP19452;
 AC

XX 26-AUG-2004 (first entry)
 DT Human secreted polypeptide #303.
 XX Human secreted polypeptide #303.
 DE Human; secreted protein; genetic disease.
 XX Homo sapiens.
 KM US2004110939-A1.
 XX 10-JUN-2004.
 PD 15-OCT-2001; 2001US-00978360.
 PF 17-DEC-1998; 98WO-1B002122.
 PR 09-FEB-1999; 99WO-1B000282.
 PR 21-JUN-2000; 2000WO-1B000951.
 PR 15-SEP-2000; 2000US-00663600.
 XX (GSEST) GENSET SA.
 PA Dumas Milne Edwards J, Bougueleret L, Jobert S, Clusel C;
 PI Duclet A;
 PI WPI; 2004-440404/41.
 DR N-PSDB; ADP19047.
 XX New isolated polynucleotide encoding secreted polypeptide, useful for
 PT gene therapy, or in diagnostic procedures to identify individuals having
 PT genetic diseases resulting from abnormal expression of the genes.
 PS Claim 2; SEQ ID NO 708; 113pp; English.
 XX The invention relates to human cDNA sequences that encode human secreted
 CC proteins. The invention also relates to an antibody that specifically
 CC binds to a polypeptide of the invention and a method of binding the
 CC polypeptide to an antibody. The polynucleotides are useful for expressing
 CC the entire secreted proteins which they encode and for distinguishing
 CC human tissues and cells from non-human tissues and cells, and for
 CC distinguishing between human tissues and cells that do or do not express
 CC the polynucleotides comprising the cDNAs. The polynucleotides and
 CC polypeptides are useful in forensic procedures or diagnostic procedures
 CC to identify individuals with genetic diseases or diagnostic procedures
 CC also useful in gene therapy to control or treat genetic diseases. This
 CC sequence represents a human secreted polypeptide of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html.
 XX Sequence 150 AA;
 SQ
 Query Match 73.8%; Score 62; DB 8; Length 150;
 Best Local Similarity 72.7%; Pred. No. 0.12;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HNNHNNHNNH 11
 Db 124 HSHNHSHSHNH 134
 RESULT 46
 AAG20566
 ID AAG20566 standard; protein, 283 AA.
 XX AAG20566;
 AC 17-OCT-2000 (first entry)
 DT Arabidopsis thaliana protein fragment SEQ ID NO: 22808.
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 22808.
 KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139482P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142277P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-015138P.
PR 01-SEP-1999; 99US-0151930P.

PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144311P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144864P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145122P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-01452913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147319P.
PR 05-AUG-1999; 99US-0147320P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.

PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 73.8%; Score 62; DB 3; Length 319;
Best Local Similarity 66.7%; Pred. No. 0.27;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNNNNNNNNN 12
Db 91 HNNNNNNHHHS 102

RESULT 48
ID AAG20564 standard; protein; 329 AA.
XX AAG20564;
AC AAG20564;
XX 17-OCT-2000 (first entry)
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 22806.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22806.
XX Protein identification; signal transduction pathway; metabolic pathway.
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX Arabidopsis thaliana.
XX EPI033405-A2.
PN EPI033405-A2.
XX 06-SEP-2000.
PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
PF 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130044P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132048P.
PR 05-MAY-1999; 99US-0132407P.
PR 06-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135124P.
PR 24-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0135829P.
PR 27-MAY-1999; 99US-0136021P.
PR 28-MAY-1999; 99US-0136392P.
PR 01-JUN-1999; 99US-0136782P.
PR 03-JUN-1999; 99US-0137222P.
PR 04-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137502P.
PR 08-JUN-1999; 99US-0137724P.
PR 10-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 22-JUN-1999; 99US-0139817P.
PR 23-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 28-JUN-1999; 99US-0140659P.
PR 28-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0140991P.
PR 01-JUL-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0143055P.
PR 08-JUL-1999; 99US-0143390P.
PR 09-JUL-1999; 99US-0143803P.
PR 12-JUL-1999; 99US-0142970P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144352P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145087P.
PR 23-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 26-JUL-1999; 99US-0145218P.
PR 27-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147302P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 10-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148655P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149368P.
PR 18-AUG-1999; 99US-0149175P.
PR 20-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 22-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149902P.
PR 25-AUG-1999; 99US-0149930P.
PR 26-AUG-1999; 99US-0150566P.
PR 27-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154018P.
PR 20-SEP-1999; 99US-0154039P.
PR 22-SEP-1999; 99US-0154779P.
PR 23-SEP-1999; 99US-0155139P.
PR 24-SEP-1999; 99US-0155486P.
PR 28-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.

PR	7-OCT-1999	99US-0158023P
PR	08-OCT-1999	99US-0158232P
PR	12-OCT-1999	99US-0158365P
PR	13-OCT-1999	99US-0159283P
PR	13-OCT-1999	99US-0159294P
PR	13-OCT-1999	99US-0159295P
PR	14-OCT-1999	99US-0159329P
PR	14-OCT-1999	99US-0159330P
PR	14-OCT-1999	99US-0159331P
PR	14-OCT-1999	99US-0159637P
PR	14-OCT-1999	99US-0159638P
PR	18-OCT-1999	99US-0159584P
PR	21-OCT-1999	99US-0160741P
PR	21-OCT-1999	99US-0160767P
PR	21-OCT-1999	99US-0160768P
PR	21-OCT-1999	99US-0160770P
PR	21-OCT-1999	99US-0160814P
PR	21-OCT-1999	99US-0160815P
PR	22-OCT-1999	99US-0160980P
PR	22-OCT-1999	99US-0160981P
PR	22-OCT-1999	99US-0160989P
PR	25-OCT-1999	99US-0161404P
PR	25-OCT-1999	99US-0161405P
PR	25-OCT-1999	99US-0161406P
PR	26-OCT-1999	99US-0161359P
PR	26-OCT-1999	99US-0161360P
PR	26-OCT-1999	99US-0161361P
PR	28-OCT-1999	99US-0161920P
PR	28-OCT-1999	99US-0161992P
PR	28-OCT-1999	99US-0162143P
PR	29-OCT-1999	99US-0162144P

Query Match	73.8%	Score 62;	DB 3;	Length 329;
Best Local Similarity	66.7%	Pred. No. 0.28;		
Matches	8;	Conservative	4;	Mismatches 0;
				Indels 0;
				Gaps 0;

QY	1	HNHNHNHNHNHN	12
		: : : :	
Db	101	HNHNHNHNHNHN	1122

RESULT 49
AAU93045
ID AAU93045 standard; protein, 329 AA
XX
AC AAU93045;

KM Agriculture; metabolic chemical; environmental stress; drought;
KM microbial disease resistance; herbicide resistance; seed yield;
KM fruit yield; growth rate; leaf senescence; flower senescence; plant
transcription factor; transgenic.

OS *Arabidopsis thaliana*

PN WO200215675-A1.

PD 28-FEB-2002

PF 22-AUG-2001; 2001WO-US026189.

PR 22-AUG-2000; 2000US-0227439P.

PR 18-APR-2001; 2001US-00837944.

PA (MEND-) MENDEL BIOTECHNOLOGY INC

PA	(PILG/)	PILGRIM M.
PA	(CREE/)	CREELMAN R.
PA	(DUBE/)	DUBELL A J.
PA	(HEAR/)	HEARD J.

PA (JIAN/) JIANG C.
PA (KEDD/) KEDDIE J.
PA (ADAM/) ADAM L.
PA (RATC/) RATCLIFFE O.
PA (REUB/) REUBER J L.
PA (RIEC/) RIECHMANN J L.
PA (YUGG/) YU G.
PA (PINE/) PINEDA O.

PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J
 P1 Adam L, Ratcliffe O, Reuber JL, Riechmann JL, Yu G, Pineda O;
 XX
 DR WPI; 2002-292022/33.
 DR N-PSDB; ABK5231.

AA An isolated or recombinant polynucleotide used to produce a transgenic plant.
PT

PS Claim 40; Page 369-370; 941pp; English.

CC The invention relates to 1 of 232 isolated or recombinant polynucleotides
CC encoding an Arabidopsis thaliana transcription factor, their variants,
CC complements, fragments, or related polynucleotide with 31% to 95%
CC sequence identity, where the plant possesses an altered trait as compared
CC to a wild-type or reference plant, or the plant exhibits an altered
CC phenotype as compared to a wild-type or reference plant, or the plant
CC exhibits ectopic expression or altered expression of one or more genes
CC associated with a plant trait as compared to a wild plant. Also included
CC are a transgenic plant comprising the polynucleotide, a computer
CC readable medium having stored sequence information, and identifying a
CC homologue sequence from a database comprising a plurality of known plant
CC sequences comprising inputting sequence information selected from one of
CC 464 fully defined sequences given in the specification. The isolated or
CC recombinant polynucleotide is used for producing a plant having a
CC modified trait, the method comprising selecting a polynucleotide that
CC encodes a polypeptide or an antisense nucleic acid, inserting the
CC polynucleotide or antisense nucleic acid into an expression vector,
CC introducing the vector into a plant or a cell of a plant to overexpress
CC the polypeptide or antisense nucleic acid, thereby producing a modified
CC plant, and selecting for a modified trait (e.g. increased production of
CC agriculturally useful proteins or metabolic chemicals, pest tolerance,
CC environmental stress response (e.g. drought), microbial disease
CC resistance, herbicide resistance, seed and fruit yield, growth rate, leaf
CC and flower senescence and many other traits listed in the specification).
CC The present sequence is one of the 232 proteins which are *A. thaliana*
CC transcription factors

Query Match	73.8%	Score 62;	DB 5;	Length 329;
Best Local Similarity	66.7%	Pred. NO.	0.28;	
Matches 8; Conservative	4;	Mismatches	0;	Gaps 0;

Qy	1	HNHNHNHNHNHN	12
		: : : :	
Db	101	HNHNHNHNHNHS	112

RESULT 50
ADCC46625
ID ADCC46625 standard; protein; 329 AA

DE Thalecress transcription factor-like protein G1052.

KW plant size; stress tolerance; yield; disease resistance; plant.

OS *Arabidopsis thaliana*

XX

PN US2003093837-A1.
 XX
 PD 15-MAY-2003.
 XX
 PF 01-NOV-2002; 2002US-00286264.
 XX
 PR 23-MAR-1999; 99US-0125814P.
 XX 22-MAR-2000; 2000US-00533030.
 PA (KEDD/) KEDDIE J.
 PA (RIEC/) RIECHMANN J L.
 PA (RATC/) RATCLIFFE O.
 PA (ZHAN/) ZHANG J.
 PA (JIAN/) JIANG C.
 PA (PINE/) PINEDA O.
 PA (HEAR/) HEARD J.
 PA (YUGG/) YU G.
 PA (ADAM/) ADAM L.
 PA (BROU/) BROUN P.
 PA (REUB/) REUBER L.
 PA (PILG/) PILGRIM M.
 PA (SAMA/) SAMAHA R.
 XX
 PI Keddle J, Riechmann JL, Ratcliffe O, Zhang J, Jiang C, Pineda O,
 PI Heard J, Yu G, Adam L, Broun P, Reuber L, Pilgrim M, Samaha R;
 DR WPI; 2003-765498/72.
 XX N-PSDB; ADC46624.
 PT Novel transgenic plant having recombinant polynucleotide encoding
 PT polypeptide that alters trait of transgenic plant when compared with same
 XX trait of another plant lacking recombinant polynucleotide.
 PS Claim 1; SEQ ID NO 24; 165pp; English.
 CC The invention relates to a transgenic plant having recombinant
 CC polynucleotide (II) encoding polypeptide comprising at least 6
 CC consecutive amino acids of a sequence chosen from the protein sequence
 CC appearing as ADC46603 - ADC46749 (every second sequence), where
 CC recombinant polynucleotide alters a trait of the seed transgenic plant
 CC when compared with same trait of another plant lacking recombinant
 CC polynucleotide. The proteins are transcription factor-like proteins. Also
 CC included are altering (M1) a trait associated with seed (comprising:
 CC transforming a plant with (II); selecting the transformed plants; and
 CC (M2) the expression levels of at least one gene of a plant (involving
 CC transforming the plant with (II) and selecting the transformed plant),
 CC altering (M3) a trait associated with a plant's seed (comprising:
 CC transforming the plant with a recombinant polynucleotide comprising a
 CC nucleotide sequence comprising least 18 consecutive nucleotides of a
 CC sequence appearing as ADC46750 - ADC46766 and selecting the transformed
 CC plant) altering (M4) a plant's trait (involving providing a database
 CC sequence, comparing the database sequence with a polypeptide or a
 CC polynucleotide chosen as detailed above, selecting a database sequence
 CC that needs selected sequence criteria and transforming a database
 CC sequence in the plant) and altering a plant's trait (involving providing
 CC a test polynucleotide, hybridizing the test polynucleotide with a
 CC polynucleotide detailed above at low stringency and transforming the
 CC hybridizing test polynucleotide in a plant to alter a trait of the
 CC plant). The method (M1) is useful for altering a trait associated with
 CC least one gene of a plant. The method (M2) is useful for altering a trait
 CC associated with a plant's seed. The method (M3) is useful for altering a trait
 CC plant's trait. The method (M4) is useful for altering a plant's trait
 CC such as seed or plant size, stress tolerance, yield or disease
 CC resistance. The present sequence represents a transcription factor-like
 XX protein/seed trait altering protein of the invention.
 SQ Sequence 329 AA;

Query Match 73.8%; Score 62; DB 7; Length 329;
 Best Local Similarity 66.7%; Pred. No. 0.28;
 Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
 Db 101 HNNNNNNHHHS 112
 RESULT 51
 ID ADD31049
 AC ADD31049 standard; protein; 329 AA.
 AC ADD31049;
 DT 15-JAN-2004 (first entry)
 DE Plant yield-related protein from clone G1052.
 XX
 KW transcription factor; transgenic plant; growth rate; senescence;
 XX seed germination rate; plant vigor; seedling vigor.
 OS Arabidopsis thaliana.
 XX
 PN WO2003013227-A2.
 PD 20-FEB-2003.
 XX
 PF 09-AUG-2002; 2002WO-US025805.
 XX
 PR 09-AUG-2001; 2001US-0310847P.
 PR 19-NOV-2001; 2001US-0336049P.
 PR 11-DEC-2001; 2001US-0338692P.
 PR 14-JUN-2002; 2002US-00171468.
 XX
 PA (MEND-) MENDEL BIOTECHNOLOGY INC.
 XX
 PI Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JB;
 PI Pilgrim M, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
 PI Broun PE;
 DR WPI; 2003-248221/24.
 XX N-PSDB; ADD31048.
 PT New plant transcription factor polynucleotides and polypeptides, useful
 PT in producing transgenic plants with commercially valuable properties,
 PT such as an alteration in a plant growth characteristic, e.g. growth rate
 XX or apomixis.
 PS Disclosure; SEQ ID NO 1078; 454pp; English.
 CC The invention relates to a number of isolated Arabidopsis thaliana cDNA
 CC sequences and their encoded proteins which are especially transcription
 CC factor related cDNAs and proteins. The isolated or recombinant plant
 CC transcription factor polynucleotides and polypeptides are useful in
 CC producing transgenic plants with commercially valuable properties, i.e.
 CC modified or altered desirable traits as compared to a reference plant,
 CC such as an alteration in a plant growth characteristic, e.g. growth rate,
 CC germination rate of seeds, vigor of plants and seedlings, or leaf and
 CC flower senescence. Sequence information related to the polynucleotides
 CC and polypeptides can also be used in bioinformatic search methods. The
 CC transgenic plant is useful for growing a progeny plant from a parent
 XX plant. This sequence represents one of the proteins of the invention.
 SQ Sequence 329 AA;

Query Match 73.8%; Score 62; DB 7; Length 329;
 Best Local Similarity 66.7%; Pred. No. 0.28;
 Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HNNNNNNNNNN 12
 Db 101 HNNNNNNHHHS 112
 RESULT 52

AD31461
 ID ADE31461 standard; protein; 329 AA.
 XX
 AC ADE31461;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Plant yield related protein from clone G1052.
 XX
 KM transcription factor; transgenic plant; salt stress resistance;
 KM osmotic stress resistance; freezing tolerance; drought tolerance;
 KM low humidity tolerance; radiation resistance.
 XX
 OS Arabidopsis thaliana.
 XX
 PN W02003013228-A2.
 XX
 PD 20-FEB-2003.
 XX
 PF 09-AUG-2002; 2002MO-US025808.
 XX
 PR 09-AUG-2001; 2001US-0310847P.
 PR 19-NOV-2001; 2001US-0336049P.
 PR 11-DEC-2001; 2001US-0338692P.
 PR 14-JUN-2002; 2002US-00171468.
 XX
 PA (MEND-) MENDEL BIOTECHNOLOGY INC.
 XX
 PI Heard JE, Riechmann JL, Creelman RA, Keddie J, Pilgrim ML,
 PI Dubell AN, Jiang C, Ratcliffe O, Pineda O, Yu G, Brown PE,
 DR N-PSDB; ADE31460.
 XX
 PT New plant transcription factor polynucleotides and polypeptides, useful
 PT in producing transgenic plants with commercially valuable properties,
 PT i.e. modified desirable traits, e.g. salt stress resistance or tolerance
 PT to freezing.
 XX
 PS Disclosure; SEQ ID NO 28; 311pp; English.
 XX
 CC The invention relates to a number of isolated cDNA sequences and their
 CC encoded proteins which are especially transcription factor related cDNA's
 CC and proteins. The isolated or recombinant plant transcription factor
 CC polynucleotides and polypeptides are useful in producing transgenic
 CC plants with commercially valuable properties, i.e. modified or altered
 CC desirable traits as compared to a reference plant, e.g. salt stress
 CC resistance, osmotic stress resistance, tolerance to freezing, drought,
 CC low humidity tolerance, or radiation resistance. Sequence information
 CC related to the polynucleotides and polypeptides can also be used in
 CC bioinformatic search methods. The transgenic plant is useful for growing
 CC a progeny plant from a parent plant. This sequence represents one of the
 CC proteins of the invention.
 XX
 SQ Sequence 329 AA;
 XX
 QY
 Db 1 HNNNNHHNNHN 12
 ||:||||:|:
 101 HNNNNHHNNHHS 112
 XX
 RESULT 53
 ID ADI41751
 ADI41751 standard; protein; 329 AA.
 XX
 AC ADI41751;
 XX
 DT 22-APR-2004 (first entry)
 XX

DE Plant transcription factor #107.
 XX
 KM transgenic; plant; enhanced tolerance to abiotic stress;
 KM glycosphate tolerance; hormone sensitivity; disease resistance;
 KM sugar sensing; flowering; flower structure; stem bifurcation;
 KM branching pattern; apical dominance; trichome; stem morphology;
 KM root growth; root hair; seed development; cell proliferation;
 KM cell differentiation; premature senescence; necrosis; plant size;
 KM leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
 KM plant anthocyanin; light response; shade avoidance; bioinformatic;
 KM transcription factor.
 XX
 OS Arabidopsis thaliana.
 XX
 PN US2004019927-A1.
 XX
 PD 29-JAN-2004.
 XX
 PF 25-FEB-2003; 2003US-00374780.
 XX
 PR 18-APR-2001; 2001US-00837944.
 XX
 PA (SHER/) SHERMAN B K.
 PA (RIEC/) RIECHMANN J L.
 PA (JIANG/) JIANG C.
 PA (HEAR/) HEARD J E.
 PA (HAAR/) HAAKE V.
 PA (CREE/) CREELMAN R A.
 PA (RATC/) RATCLIFFE O.
 PA (ADAM/) ADAM L J.
 PA (REUB/) REUBER T L.
 PA (KEDD/) KEDDIE J.
 PA (BROU/) BROUN P E.
 PA (PIIG/) PILGRIM M L.
 PA (DUBE/) DUBELL A N.
 PA (PINE/) PINEDA O.
 PA (YUGG/) YU G.
 XX
 PI Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V,
 PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Brown PE,
 PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
 DR N-PSDB; ADI41750.
 XX
 PT New transgenic plant comprising a recombinant polynucleotide of any one
 PT of more than 500 nucleotide sequences, useful in bioinformatic search
 PT methods.
 XX
 PS Claim 1; SEQ ID NO 214; 435pp; English.
 XX
 CC The invention describes a transgenic plant comprising a recombinant
 CC polynucleotide of any one of more than 500 nucleotide sequences fully
 CC defined in the specification or its complement. The method of the
 CC invention can be used to produce a plant having altered traits such as:
 CC enhanced tolerance to abiotic stress; glycosphate tolerance; hormone
 CC sensitivity; disease resistance; sugar sensing; early or late flowering;
 CC altered flower structure; change in stem bifurcations; altered branching
 CC pattern; reduced apical dominance; reduced trichome density; lack of
 CC trichomes; reduced ectopic trichome development; altered trichome
 CC development; increase in trichome number; altered stem morphology;
 CC increased root growth; increased root hairs; altered seed development;
 CC altered cell proliferation or cell differentiation; rapid development;
 CC premature senescence; increased necrosis; increase in seedling or plant
 CC size; decreased plant size; leaf morphology; seed morphology; seed
 CC biochemistry; increase in root anthocyanins; increase in plant
 CC anthocyanins; or alternation in light response or shade avoidance. The
 CC transgenic plant, polynucleotides and polypeptides are useful in
 CC bioinformatic search methods. This is the amino acid sequence of a plant
 CC transcription factor that can be used in the creation of a transgenic
 CC plant with altered traits.
 XX
 SQ Sequence 329 AA;
 XX

Query Match 73.8%; Score 62; DB 8; Length 329;
 Best Local Similarity 66.7%; Pred. No. 0.28;
 Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 HNNNNNNNNNN 12
 Db 101 HNNNNHHHHHS 112

RESULT 54
 ADI61339
 ID ADI61339 standard; protein: 329 AA.
 AC ADI61339;
 DT 22-APR-2004 (first entry)
 XX
 DE A. thaliana novel transcription factor (TF) #19.
 XX
 KM Plant; transcription factor; TF; plant trait.
 XX
 OS Arabidopsis thaliana.
 XX
 PN US2003229915-A1.
 PD 11-DEC-2003.
 XX
 PF 22-NOV-2002; 2002US-00302267.
 XX
 PR 18-FEB-1999; 99US-0120880P.
 PR 22-FEB-1999; 99US-0121037P.
 PR 11-MAR-1999; 99US-0124276P.
 PR 15-APR-1999; 99US-0129450P.
 PR 20-MAY-1999; 99US-0135134P.
 PR 15-JUL-1999; 99US-0144153P.
 PR 22-OCT-1999; 99US-0161143P.
 PR 01-NOV-1999; 99US-0162656P.
 PR 17-FEB-2000; 2000US-00506720.
 XX
 PA (KEDD/) KEDDIE J.
 PA (FROM/) FROMM M.
 PA (HEAR/) HEARD J.
 PA (RIEC/) RIECHMANN J L.
 PA (ADAM/) ADAM L.
 PA (BROU/) BROUN P.
 PA (PINE/) PINEDA O.
 PA (REUB/) REUBER L.
 PA (ZHAN/) ZHANG J.
 PA (YUGG/) YU G.
 PA (JIANG/) JIANG C.
 PA (SAMA/) SAMAH R.
 PA (PIIG/) PILGRIM M.
 PA (CREE/) CREELMAN R.
 PI Keddie J, Fromm M, Heard J, Riechmann JL, Adam L, Broun P,
 PI Pineda O, Reuber L, Zhang J, Yu G, Jiang C, Samaha R, Pilgrim M,
 PI Creelman R;
 DR WPI: 2004-052052/05.
 DR N-PDB; ADI61338.
 PT New polynucleotide, useful in screening for a transcription factor that
 PT modifies a plant trait.
 XX
 PS Claim 9; SEQ ID NO 38; 16bp; English.
 XX
 CC The present invention relates to the isolation of novel plant
 CC (arabidopsis thaliana) polynucleotide sequences that encode transcription
 CC factors (TFs), and the polypeptide sequences for the TFs. The
 CC polynucleotide sequences are useful in screening for a transcription
 CC factor that modifies a plant trait. Also disclosed is an expression
 CC vector comprising a TF polynucleotide sequence, a host cell comprising

CC the expression vector, a transgenic plant comprising or ectopically
 CC expressing an isolated TF polynucleotide sequence, a method for screening
 CC for a molecule that modifies a plant trait, a method for producing a TF
 CC transgenic plant, a method for identifying a sequence homologous to a TF
 CC polynucleotide or polypeptide sequence, and a method for screening for a
 CC transcription factor that modifies a plant trait. The present sequence
 CC represents a novel A. thaliana TF of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification. The
 CC complete sequence data for this patent was obtained in electronic format
 CC directly from the USPTO web site at seqdata.uspto.gov.
 XX
 SO Sequence 329 AA;

Query Match 73.8%; Score 62; DB 8; Length 329;
 Best Local Similarity 66.7%; Pred. No. 0.28;
 Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 HNNNNNNNNNN 12
 Db 101 HNNNNHHHHHS 112

RESULT 55
 ADO02287
 ID ADO02287 standard; protein: 329 AA.
 AC ADO02287;
 DT 01-JUL-2004 (first entry)
 XX
 DE Thalecress transcription factor protein #350.
 XX
 KM Thalecress; transcription factor; plant; transgenic; abiotic stress;
 KM cold tolerance; heat tolerance; drought; osmotic stress;
 KM phosphate limitation; disease resistance; sugar sensing; seed germination;
 KM hormone sensitivity; inflorescence architectural change;
 KM flowering; inflorescence architectural change;
 KM meristem cell differentiation; phylloclax; apical dominance;
 KM trichome development; seed development; premature senescence;
 KM delayed senescence; lethality; necrosis; plant size; leaf morphology;
 KM seed morphology; secondary metabolism; light response; shade avoidance.
 XX
 OS Arabidopsis thaliana.
 XX
 PN US2004045049-A1.
 PD 04-MAR-2004.
 XX
 PF 10-APR-2003; 2003US-00412699.
 XX
 PR 13-SEP-1999; 99US-00394519.
 PR 21-JAN-2000; 2000US-00489376.
 PR 17-FEB-2000; 2000US-00506720.
 PR 22-MAR-2000; 2000US-00532591.
 PR 22-MAR-2000; 2000US-00533029.
 PR 22-MAR-2000; 2000US-00533030.
 PR 22-MAR-2000; 2000US-00533392.
 PR 22-MAR-2000; 2000US-00533648.
 PR 06-APR-2000; 2000US-00509448.
 PR 16-NOV-2000; 2000US-00713994.
 PR 27-MAR-2001; 2001US-00819142.
 PR 17-APR-2001; 2001US-00837444.
 PR 30-JAN-2002; 2002US-00958131.
 PR 14-JUN-2002; 2002US-00171468.
 PR 09-AUG-2002; 2002US-00225066.
 PR 09-AUG-2002; 2002US-00225067.
 PR 09-AUG-2002; 2002US-00225067.
 PR 17-DEC-2002; 2002US-0434166P.
 PR 25-FEB-2003; 2003US-00374780.
 XX
 PA (ZHAN/) ZHANG J.
 PA (FROM/) FROMM M E.
 PA (HEAR/) HEARD J E.

PA (RIEC/) RIECHMANN J L.
 PA (ADAM/) ADAM L J.
 PA (BROU/) BROUN P E.
 PA (PINE/) PINEDA O.
 PA (REUB/) REUBER T L.
 PA (KEDD/) KEDDIE J S.
 PA (YUGG/) YU G.
 PA (JIANG/) JIANG C.
 PA (SAMAH/) SAMAHA R S.
 PA (PIILG/) PILGRIM M L.
 PA (CREE/) CREELMAN R A.
 PA (DUBE/) DUBELL A N.
 PA (RATC/) RATCLIFFE O.
 PA (KUMI/) KUMIMOTO R.
 PA (SHER/) SHERMAN B K.
 XX
 PI Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE,
 PI Pineda O, Reuber TL, Keddie JS, Yu G, Jiang C, Samaha RS,
 PI Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R,
 PI Sherman BK;
 XX
 DR WPI; 2004-225755/21.
 DR N-PSDB; AD002286.
 XX
 PT New transgenic plant, useful in developing phenotypes with altered or
 PT improved characteristics or traits.
 XX
 PS Claim 1; SEQ ID NO 700; 213pp; English.
 XX
 CC The invention relates to a transgenic plant comprising a recombinant
 CC polynucleotide having a polynucleotide sequence or its complementary
 CC sequence comprising a sequence encoding a polypeptide, that initiates
 CC transcription (i.e. a transcription factor) from Arabidopsis, Soybean,
 CC Rice, Rape or Corn, comprising any of the sequences appearing as AD001588
 CC -AD003527 or AD003530-AD003559. Also included are using a transgenic
 CC plant to grow a progeny plant, an expression cassette (comprising a
 CC constitutive, inducible or tissue-specific promoter and a recombinant
 CC polynucleotide described above), a host cell comprising the expression
 CC cassette, producing a modified plant having a modified trait, identifying
 CC a factor that is modulated by or interacts with a polypeptide encoded by
 CC the polynucleotide sequence and identifying at least one downstream
 CC polynucleotide sequence that is subject to a regulatory effect of any of
 CC the polypeptides encoded by the polynucleotide described above. The
 CC transgenic plant is useful for producing a plant that has an altered
 CC trait e.g. an enhanced tolerance to abiotic stresses (increased tolerance
 CC to chilling, germination in cold conditions, freezing tolerance, tolerance
 CC to heat, tolerance to drought, tolerance to osmotic stress, tolerance to
 CC salt, tolerance to phosphate limitation, tolerance to potassium
 CC limitation, decreased sensitivity to nitrogen limitation), altered
 CC hormone sensitivity, reduced sensitivity to abscisic acid, an altered
 CC response to ethylene, disease resistance, altered susceptibility to
 CC Botrytis, altered susceptibility to Fusarium, altered susceptibility to
 CC Erysiphe, altered susceptibility to Pseudomonas syringae, altered
 CC susceptibility to Sclerotinia, altered sugar sensing, improved seed
 CC germination and seedling vigor, early flowering, late flowering, extended
 CC period of flowering, an inflorescence architectural change, a change in
 CC stem bifurcations, a lack of a shoot meristem, reduced meristem cell
 CC differentiation, altered phyllotaxy, altered branching pattern, reduced
 CC apical dominance, reduced trichome density, ectopic trichome development,
 CC altered trichome development, altered stem morphology, increased root
 CC growth, increased root hairs, altered seed development, altered cell
 CC proliferation/cell differentiation, premature senescence, delayed
 CC senescence, lethality, increased necrosis, an increase in seedling or
 CC plant size, decreased plant size, a change in leaf morphology, increased
 CC altered leaf development, increased leaf size and mass, glossy leaves,
 CC leaf cell expansion, change in seed morphology, altered seed coloration,
 CC increased seed size, decreased seed size, altered seed shape, change in
 CC leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid
 CC content, increased leaf insoluble sugars, decreased leaf insoluble
 CC sugars, increased leaf anthocyanins, an alteration of leaf fatty acid
 CC content, an alteration of leaf glucosinolate content, change in seed
 CC biochemistry, an increase in seed oil content, decrease in seed oil
 CC content, increase in seed fatty acid content, decrease in seed fatty acid

CC content, increase in seed protein content, decrease in seed protein
 CC content, alteration in seed prenyl lipid content, increase in seed
 CC sterols, upregulation of genes involved in secondary metabolism, increase
 CC in root anthocyanins, increase in plant anthocyanins, and alteration in
 CC light response or shade avoidance. The present sequence represents a
 CC chalcress transcription factor of the invention.
 XX
 SQ : Sequence 329 AA;
 Query Match 73.8%; Score 62; DB 8; Length 329;
 Best Local Similarity 66.7%; Pred. No. 0.28;
 Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 HHHHHHHHHHHN 12
 ||:||||:|:
 Db 101 HHHHHHHHHHS 112
 RESULT 56
 ABP41168
 ID ABP41168 standard; protein; 476 AA.
 XX
 AC ABP41168;
 XX
 DT 23-AUG-2002 (first entry)
 XX
 DB Human ovarian antigen HDPF41, SEQ ID NO:2300.
 XX
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; reproductive system disorder; infertility;
 KW pregnancy disorder; anovulation; polycystic ovary syndrome; PCOS;
 KW ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive.
 XX
 OS Homo sapiens.
 XX
 PN MO200200677-A1.
 XX
 PD 03-JAN-2002.
 XX
 PF 07-JUN-2001; 2001MO-US018569.
 XX
 PR 07-JUN-2000; 2000US-0209467P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI; 2002-147878/19.
 DR N-PSDB; ABC54245.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
 PT cancer), immune disorders, cardiovascular disorders and neurological
 PT diseases.
 XX
 PS Claim 11; SEQ ID NO 2300; 2922pp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (AB054131-AB056305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system

CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders,
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence data represents a human ovarian antigen of the invention. Note: The
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX

Sequence 476 AA:

Query Match Best Local Similarity 73.8%; Score 62; DB 5; Length 476;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
|:|:|:|:|:|:|:
Db 133 HSHNHQSHNH 143

RESULT 57
AB61643
ID AB61643 standard; protein; 601 AA.
XX
XX AB61643;
AC

DT 26-MAR-2002 (first entry)
XX

DE Drosophila melanogaster polypeptide SEQ ID NO 11721.
XX

KW Drosophila; developmental biology; cell signalling; insecticide;
XX
XX pharmaceutical.
KM

OS Drosophila melanogaster.
XX

PN WO200171042-A2.
XX

PD 27-SEP-2001.
XX

PF 23-MAR-2001; 2001WO-US009231.
XX

PR 23-MAR-2000; 2000US-0191637P.
XX

PT 11-JUL-2000; 2000US-00614150.
XX

PA (PEKE) PE CORP NY.
XX

PI Venter JC, Adams M, Li PMD, Myers EW,
XX

DR WPI; 2001-656860/75.
XX

DR N-PSDB; ABL05746.
XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX

PS Disclosure; SEQ ID NO 11721; 21pp + Sequence listing; English.
XX

CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-AB130511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (AB65737-
XX AB672072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 601 AA:

Query Match Best Local Similarity 73.8%; Score 62; DB 4; Length 601;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:|:|:
Db 250 HHHHHHNNNN 261

RESULT 58
AAU74629
ID AAU74629 standard; protein; 829 AA.
XX

AC AAU74629;
XX

DT 09-APR-2002 (first entry)
XX

DE Oestrogen-regulated LIV-1 family protein BAA6579_Hs.
XX

KW LIV-1; oestrogen; cytosolic; neuroprotective; zinc homeostasis;
XX gene therapy; apoptosis modulator; cancer; neurodegenerative disorder;
XX apoptotic disorder; zinc-homeostasis related disorder.
XX

OS Homo sapiens.
XX

PN WO200196372-A2.
XX

PD 20-DEC-2001.
XX

PF 13-JUN-2001; 2001WO-GB002597.
XX

PR 13-JUN-2000; 2000GB-00014411.
XX

PR 14-JUN-2000; 2000GB-00014493.
XX

PR 05-JUL-2000; 2000US-0216349P.
XX

PA (UYCA-) UNIV COLLEGE CARDIFF.
XX

PI Taylor KM, Morgan HE, Nicholson RI;
XX

DR WPI; 2002-106465/14.
XX

PT Use of a polypeptide comprising one or more consensus regions of proteins
XX of LIV-1 family for treating disorders of zinc homeostasis such as breast
XX cancer, neurodegenerative disorders, and for modifying apoptosis.
XX

PS Claim 1; Fig 1; 67pp; English.
XX

CC The invention describes the a polypeptide comprising one or more
XX consensus regions of proteins of LIV-1 family or its functional
XX homologue. The polypeptide is useful in the preparation of a medicament
XX for the treating a disease e.g. those involving disorders of zinc
XX homeostasis, in gene therapy and for modifying apoptosis in vitro or in
XX vivo on contact with cells. Diseases involving defects in zinc
XX homeostasis include cancer, neurodegenerative disorders and apoptotic
XX disorders. Recombinant proteins of the LIV-1 family (an oestrogen-
XX regulated gene) are useful for diagnosing zinc homeostasis-related
XX condition in a subject which involves contacting a sample from the
XX subject with the recombinant protein and measuring the binding of
XX antibody to the sample. The antibody is also useful for treating a zinc
XX homeostasis-related condition. This sequence is a member of the LIV-1
XX family (a gene regulated by oestrogen levels) and is useful for creating
XX recombinant proteins for diagnosing zinc-homeostasis related conditions,
XX described in the method of the invention
SQ Sequence 829 AA;

Query Match 73.8%; Score 62; DB 5; Length 829;
 Best Local Similarity 72.7%; Pred. No. 0.75;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
 | : | | | : | | |
 Db 123 HSHNHQSHNH 133

RESULT 59
 AAB87345
 ID AAB87345 standard; protein; 831 AA.
 AC AAB87345;
 XX
 DT 22-MAY-2001 (first entry)
 XX
 DE Human gene 4 encoded secreted protein HDPFY41, SEQ ID NO:86.
 XX
 KW Human; secreted protein; proliferative disorder; cancer; tumour;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; aschma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vulnetary; cell culture;
 KW chemotaxis; food additive; binding partner identification.
 XX
 OS Homo sapiens.
 XX
 PN MO200116022-A1.
 XX
 PD 15-MAR-2001.
 XX
 PF 31-AUG-2000; 2000MO-US024008.
 XX
 PR 03-SRP-1999; 99US-0152315P.
 XX
 PR 03-SRP-1999; 99US-0152317P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ni J, Baker KP, Birse CE, Fiacella M, Komatsoulis GA, Rosen CA,
 PI Soppet DR, Young PE, Edner R, Duan DR, Olsen HS, Lafleur DW;
 PI Moore PA, Shi Y, Wei Y, Florence KA;
 DR WPI; 2001-203081/20.
 DR N-PSDB; AAF91861.
 XX
 PT Nucleic acid molecules encoding human secreted proteins, used in
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 PT Parkinson's diseases and cancers.
 XX
 Claim 11; Page 533-535; 607pp; English.

AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted
 CC protein genes, and AAB87342-AAB87413 represent the proteins they encode.
 CC AAB87414-AAB87454 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new gene. Specific uses are described for each of the
 CC 52 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin
 CC disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,

CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein of the invention
 XX

Sequence 831 AA;
 SQ

Query Match 73.8%; Score 62; DB 4; Length 831;
 Best Local Similarity 72.7%; Pred. No. 0.75;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
 | : | | | : | | |
 Db 124 HSHNHQSHNH 134

RESULT 60
 ABG65404
 ID ABG65404 standard; protein; 831 AA.
 AC ABG65404;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Human albumin fusion protein #2079.
 XX
 KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
 KW human serum albumin; HSA; cancer; reproductive disorder;
 KW digestive disorder; immune disorder; endocrine disorder;
 KW haematopoietic disorder; neural disorder; connective disorder;
 KW cytostatic; antifertility; antiinflammatory; antitumor;
 KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
 KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
 KW osteopathic; antiarthritic.
 XX
 OS Homo sapiens.
 XX
 OS Synthetic.
 XX
 PN MO200177137-A1.
 XX
 PD 18-OCT-2001.
 XX
 PF 12-APR-2001; 2001MO-US011988.
 XX
 PF 12-APR-2000; 2000US-0229358P.
 XX
 PR 25-APR-2000; 2000US-0199384P.
 XX
 PR 21-DEC-2000; 2000US-0256931P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Haseltine WA;
 PI WPI; 2002-010886/01.
 XX
 PT New fusion protein for treating disease e.g. diabetes comprises an
 PT albumin fused to a therapeutic protein.
 XX
 Claim 1; Page 1972-1975; 2102pp; English.

The present invention relates to albumin fusion proteins comprising a
 CC therapeutic protein X and human albumin (HA, also known as human serum
 CC albumin, HSA). The proteins are useful for treating a disease or disorder
 CC that may be modulated by therapeutic protein X. The albumin extends the
 CC shelf-life of protein X, and may increase its biological in vitro/in vivo

CC activity. The protein is useful for treating and diagnosing disorders
 CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
 CC disease, ulcerative colitis), immune disorders (e.g. acquired
 CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
 CC hematopoietic disorders, neural disorders (e.g. Alzheimer's
 CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,
 CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).
 CC AB63326-AB65518 represent albumin fusion proteins of the invention
 XX

XX Sequence 831 AA;

Query Match

Best Local Similarity 73.8%; Score 62; DB 5; Length 831;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 HNNHNNHNNH 11
 124 HSHNHQSHNH 134

Db 124 HSHNHQSHNH 134

RESULT 61

ID ABB97347 standard; protein; 831 AA.

AC ABB97347;

DT 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 615.

KW Human; antineoplastic; vulnery; antiinflammatory; immunomodulator;
 KW antiinfectivity; cerebroprotective; cytostatic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 XX expressed sequence tag.

OS Homo sapiens.

PN WO200222660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001WO-US026015.

PR 11-SEP-2000; 2000US-00659671.

PA (HYSE-) HYSEQ INC.

FI Tang YT, Liu C, Zhou P, Agundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX WPI; 2002-292408/33.

DR N-PSDB; ABN32533.

PT An isolated polynucleotide for treating diseases associated with its
 XX encoded polypeptide such as cancer and multiple sclerosis.

PS Example 2; SEQ ID NO 615; 509pp; English.

CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention
 XX

XX Sequence 831 AA;

Query Match 73.8%; Score 62; DB 5; Length 831;
 Best Local Similarity 72.7%; Pred. No. 0.75;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 HNNHNNHNNH 11
 124 HSHNHQSHNH 134

RESULT 62

ID ADL78671 standard; protein; 831 AA.

AC ADL78671;

DT 20-MAY-2004 (first entry)

DE Albumin fusion protein related therapeutic protein X, SEQ ID NO 2153.

KW albumin fusion protein; cytostatic; antineoplastic; antiarthritis;
 KW antiallergic; anti-HIV; immunosuppressive; antiinflammatory;
 KW antiparasitic; antibacterial; osteopathic; dermatological; antigout;
 KW immunomodulator; antiarhythmic; cardiac; neurologic; antipneumatic;
 KW nephrotropic; uropathic; neuroprotective; antiparkinsonian; tranquilizer;
 KW antidiabetic; anabolic; hypertensive; vulnery; gene therapy; cancer;
 KW reproductive system disorder; therapeutic protein.

XX Unidentified.

PN US2004010134-A1.

PD 15-JAN-2004.

PF 12-APR-2001; 2001US-00833245.

PR 12-APR-2000; 2000US-0229358P.

PR 25-APR-2000; 2000US-0199384P.

PR 21-DEC-2000; 2000US-0256931P.

PA (ROSE/) ROSEN C A.

PI (HASE/) HASELTINE W A.

DR Rosen CA, Haseltine WA;

PT WPI; 2004-090519/09.

PS Disclosure; SEQ ID NO 2153; 279pp; English.

CC The invention relates to a novel albumin fusion protein. The invention
 CC further relates to: a composition comprising the albumin fusion protein
 CC and a pharmaceutical carrier; a kit comprising the composition of the
 CC albumin fusion protein formula; a method of treating a disease or
 CC disorder in a patient comprising the step of administering the albumin
 CC fusion protein; a method of treating a patient with a disease or disorder
 CC that is modulated by Therapeutic protein: X, or its fragment or variant;
 CC a method of extending the shelf life of Therapeutic protein: X, or its
 CC fragment or variant; a nucleic acid molecule comprising a polynucleotide
 CC nucleic acid encoding the albumin fusion protein; a vector comprising the
 CC nucleic acid molecule of the albumin fusion protein; and a host cell
 CC comprising the nucleic acid molecule of the albumin fusion protein. The
 CC albumin fusion protein and its compositions have the following
 CC activities: cytostatic, antineoplastic, antiarthritis, antiallergic, anti-
 CC HIV, immunosuppressive, antineoplastic, antiarthritis, antiallergic, anti-
 CC osteopathic, dermatological, antigout, immunomodulator, antipneumatic,
 CC cardiac, neurologic, antipneumatic, nephrotropic, uropathic,
 CC neuroprotective, antiparkinsonian, tranquilizer, antidiabetic, anabolic,
 CC hypertensive, and vulnery. The albumin fusion protein nucleic acid may
 CC be used in gene therapy to treat disorders. The albumin fusion protein is
 CC useful for diagnosing, treating, preventing or ameliorating diseases or
 CC disorders comprising indication: Y. The diseases or disorders include:
 CC cancer (e.g. leukemia, colon, bone, breast, liver or lung cancer).

immune or haematopoietic diseases (e.g. anaemia, Hodgkin's disease, acute lymphocytic anaemia, multiple myeloma, arthritis, asthma, AIDS, autoimmune diseases, inflammatory bowel disease, psoriasis or Lyme disease), reproductive system disorders (e.g. prostatitis, inguinal hernia, varicocele, penile carcinoma, ovarian adenocarcinoma or Sertoli-Leydig tumours), musculoskeletal diseases (e.g. giant cell tumours, Paget's disease, myelocytic lupus erythematosus, gout, muscular dystrophy or cachexia), cardiovascular disease (e.g. rhabdomyoma, heart disease, arrhythmia, cardiac arrest, heart valve disease, hypernatraemia or hyponatraemia), mixed foetal diseases (e.g. foetal alcohol syndrome, Down's syndrome, Patou syndrome, Turner's syndrome, Apert syndrome or Tay-Sachs disease), excretory diseases (e.g. urinary incontinence, urinary tract infections or renal disorders), neural or sensory disease (e.g. Alzheimer's disease, Parkinson's disease, cerebral malaria, meningitis, cerebellar ataxia, attention deficit disorder, autism or obsessive compulsive disorder), respiratory disease (e.g. emphysema, lung cancer or occupational lung disease), endocrine diseases (e.g. diabetes, Addison's disease or glomerulonephritis), digestive diseases (e.g. portal hypertension, irritable bowel disease, gastric atrophy or pancreatitis) or connective tissue or epithelial diseases (e.g. Crohn's disease, scleroderma, wound healing or epidermolysis bullosa). This sequence represents a therapeutic protein X relating to the albumin fusion protein of the invention. The sequence listing data for this specification was downloaded from the USPTO website.

SQ Sequence 831 AA;

Query Match	73.8%	Score 62;	DB 8;	Length 831;
Best Local Similarity	72.7%	Pred. NO. 0.75;		
Matches	8;	Conservative	2;	Mismatches 1;
				Indels 0;
				Gaps 0;

QY	1	HNHNHNHNHNH	1 1
		: :	
Db	1 2 4	HSNHNQHSNHNH	1 3 4

RESULT 63

ID ADQ21341 standard; protein; 835 AA

AC ADQ21341

DT 26-AUG-2004 (first entry)

DE Human soft tissue sarcoma-upregulated protein - SEQ ID 4161.

KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human

05 Homo sapiens

PN WO2004048938-A2.

PD 10-JUN-2004.

PF 26-NOV-2003; 2003WO-US038193.

PR 26-NOV-2002; 2002US-0429739P.

PA (PROT-) PROTEIN DESIGN LABS INC.

PI Aziz N, Ginsburg WM, Zlotnik A;

DR WPI; 2004-441208/41.

PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.

PS Example 2; SEQ ID NO 4161; 210pp; English.

CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual

CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cyrostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.

SQ Sequence 835 AA;

Query Match	73.8%	Score 62	DB 8	Length 835
Best Local Similarity	72.7%	Pred. NC. 0.76		
Matches	8	Conservative	2	Mismatches 1
				Indels 0
				Gaps 0

```
QY      1 HNHNNHNNHNNH 11
        |:|||:|||
Db      128 HSHNHQSHNNH 138
```

RESULT 64

ID ADQ59370 standard; protein; 883 AA

AC ADQ59370;

DT 07-OCT-2004 (first entry)

DE Human cancer-associated (CA) protein sequence SEQ ID NO:6.

KM human; cancer-associated gene; cancer-associated protein; cytostatic;
 KM gene therapy; vaccine; tyrosine kinase antagonist;
 KM G-protein coupled receptor antagonist; cancer; lymphoma.

Os "Homo sapiens"

PN WO2004058288-A1

PD 15-JUL-2004.

PF 15-DEC-2003; 2003WO-US040082.

PR 17-DEC-2002; 2002US-00322696.

PA (SAGR-) SAGRES DISCOVERY INC.

PI Morris DW, Malandro MS,

DR WPI; 2004-543349/52.

XX

PT New cancer-associated nucleic acid for diagnosing, preventing or treating
PT cancer (e.g. lymphoma) or for screening agents that may be used for
PT treating or preventing cancer.

PS Claim 19; SEQ ID NO 6; 143pp; English.

CC The present invention describes human cancer-associated (CA) nucleotide
CC sequences (1). Also described: (1) an expression vector comprising (1);
CC (2) a host cell comprising (1) or the expression vector; (3) a microarray
CC for detecting a CA nucleic acid; (4) an isolated polypeptide encoded
CC within an open reading frame of a CA sequence; (5) an isolated antibody,
CC or its antigen binding fragment, that binds to the above polypeptide; (6)
CC a hybridoma that produces the monoclonal antibody described above; (7) a
CC pharmaceutical composition comprising the antibody and a pharmaceutical
CC excipient; (8) a kit for detecting or diagnosing cancer cells, comprising
CC the above (monoclonal) antibody or polynucleotide that selectively
CC hybridizes to any of the polynucleotide sequences mentioned above; (9)
CC methods for diagnosing cancer or for detecting the presence or absence of
CC cancer cells in an individual; (10) a method for inhibiting growth of

CC cancer cells in an individual; (11) a method for delivering a therapeutic
CC agent to cancer cells in an individual; (12) an electronic library
CC comprising the polynucleotide or polypeptide, or their fragments,
CC mentioned above; (13) a method of screening for anticancer activity; (14)
CC or the presence of the antibody in a test cell or serum sample; (15) a
CC method for screening for a bioactive agent capable of modulating the
CC activity of a CA protein encoded by the above nucleic acid molecule; and
CC (16) a method for treating cancers. (1) has cytoskeletal activity; and can
CC be used in gene therapy, in vaccines, as a tyrosine kinase antagonist,
CC and as a G-protein coupled receptor antagonist. The compositions and
CC methods of the present invention can be used for diagnosing, preventing
CC and treating cancer, especially lymphomas. They may also be used in
CC screening for agents that may be used for treating or preventing cancer.
CC The present sequence represents a human CA gene amino acid sequence.
CC which is given in the exemplification of the present invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 883 AA;

XX SQ

XX Query Match

XX Best Local Similarity 73.8%; Score 62; DB 8; Length 883;

XX Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 1 HNNHNNHNNH 11
176 HSHNHQSHNH 186

XX RESULT 65

XX ABU24645

XX ID ABU24645 standard; protein; 429 AA.

XX AC ABU24645;

XX XX

XX DT 19-JUN-2003 (first entry)

XX DE

XX Proteins encoded by Prokaryotic essential gene #10172.

XX KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Clostridium botulinum.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA28515.

XX XX

XX PT New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 52569; 1766bp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of

XX CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC product's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 429 AA;

XX Query Match

XX Best Local Similarity 72.6%; Score 61; DB 6; Length 429;

XX Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 1 HNNHNNHNNH 12
92 HNHSHSHHNN 103

XX RESULT 66

XX ABB65764

XX ID ABB65764 standard; protein; 850 AA.

XX AC ABB65764;

XX XX

XX DT 26-MAR-2002 (first entry)

XX DE

XX Drosophila melanogaster polypeptide SEQ ID NO 24084.

XX KM Drosophila; developmental biology; cell signalling; insecticide;

XX KM pharmacological.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX PI WPI; 2001-656860/75.

XX DR N-PSDB; ABL09867.

XX XX

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 24084; 21pp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB160511), expressed DNA
 CC sequences (AB161840-AB16175) and the encoded proteins (AB57737-
 CC AB572072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 850 AA;
 Query Match 72.6%; Score 61; DB 4; Length 850;
 Best Local Similarity 66.7%; Pred. No. 1.1;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HNNNNNNNNNN 12
 |||:|:|:|:
 144 HNNHHHHOHN 155
 DB
 RESULT 67
 ABB62708
 ID ABB62708 standard; protein; 1028 AA.
 XX
 AC ABB62708;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE *Drosophila melanogaster* polypeptide SEQ ID NO 14916.
 XX
 KM *Drosophila*; developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 XX
 OS *Drosophila melanogaster*.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; AB1606811.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 14916; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB160511), expressed DNA
 CC sequences (AB161840-AB16175) and the encoded proteins (AB57737-
 CC AB572072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1028 AA;
 Query Match 72.6%; Score 61; DB 4; Length 1028;
 Best Local Similarity 66.7%; Pred. No. 1.3;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HNNNNNNNNNN 12
 |||:|:|:|:
 144 HNNHHHHOHN 155
 DB
 RESULT 68
 ABG05885
 ID ABG05885 standard; protein; 59 AA.
 XX
 AC ABG05885;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #5876.
 XX
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 XX
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Dmanac RT, Liu C, Tang YF;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS70072.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 36244; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG030377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

Sequence 59 AA;
Query Match 70.2%; Score 59; DB 4; Length 59;
Best Local Similarity 58.3%; Pred. No. 0.12;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HHHHHHHHHN 12
Db 20 HHHHHHHHHN 31
RESULT 69
AA616190
ID AA616190 standard; protein; 83 AA.
XX
AC AA616190;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 16735.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
OS
XX EPI033405-A2.
FN
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-00301439.
PF
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 07-MAY-1999; 99US-0132486P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135553P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137223P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137722P.
PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 28-JUN-1999; 99US-0140635P.
PR 29-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0140991P.
PR 01-JUL-1999; 99US-0141287P.
PR 02-JUL-1999; 99US-0141842P.
PR 06-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142350P.
PR 09-JUL-1999; 99US-0142803P.
PR 12-JUL-1999; 99US-0142920P.
PR 13-JUL-1999; 99US-0142977P.
PR 14-JUL-1999; 99US-0143542P.
PR 15-JUL-1999; 99US-0143624P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144633P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 22-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 27-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147326P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.

PR	09-AUG-1999;	99US--0147931P.
PR	10-AUG-1999;	99US-0146171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149366P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149320P.
PR	23-AUG-1999;	99US-0149320P.
PR	23-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157533P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160982P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161931P.
PR	29-OCT-1999;	99US-0162142P.

Db	55	NNHSHNONHSHN	66
RESULT 70			
AAAG16189			
ID	AAAG16189	standard; protein; 120 AA.	
XX			
AC	AAAG16189;		
XX			
DT	17-OCT-2000	(first entry)	
XX			
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 16734.		
XX			
KW	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-00301439.		
XX			
25-FEB-1999;	99US-0121825P.		
PR	05-MAR-1999;	99US-0123180P.	
PR	09-MAR-1999;	99US-0123548P.	
PR	23-MAR-1999;	99US-0125788P.	
PR	25-MAR-1999;	99US-0126284P.	
PR	29-MAR-1999;	99US-0126785P.	
PR	01-APR-1999;	99US-0127462P.	
PR	06-APR-1999;	99US-0128234P.	
PR	08-APR-1999;	99US-0128714P.	
PR	16-APR-1999;	99US-0129845P.	
PR	19-APR-1999;	99US-0130077P.	
PR	21-APR-1999;	99US-0130449P.	
PR	23-APR-1999;	99US-0130510P.	
PR	23-APR-1999;	99US-0130891P.	
PR	28-APR-1999;	99US-0131449P.	
PR	30-APR-1999;	99US-0132048P.	
PR	30-APR-1999;	99US-0132407P.	
PR	04-MAY-1999;	99US-0132484P.	
PR	05-MAY-1999;	99US-0132485P.	
PR	06-MAY-1999;	99US-0132486P.	
PR	06-MAY-1999;	99US-0132487P.	
PR	07-MAY-1999;	99US-0132863P.	
PR	11-MAY-1999;	99US-0134263P.	
PR	14-MAY-1999;	99US-0134218P.	
PR	14-MAY-1999;	99US-0134219P.	
PR	14-MAY-1999;	99US-0134221P.	
PR	14-MAY-1999;	99US-0134570P.	
PR	18-MAY-1999;	99US-0134768P.	
PR	19-MAY-1999;	99US-0134941P.	
PR	20-MAY-1999;	99US-0135124P.	
PR	21-MAY-1999;	99US-0135353P.	
PR	24-MAY-1999;	99US-0135629P.	
PR	25-MAY-1999;	99US-0136021P.	
PR	27-MAY-1999;	99US-0136392P.	
PR	28-MAY-1999;	99US-0136782P.	
PR	01-JUN-1999;	99US-0137222P.	
PR	03-JUN-1999;	99US-0137528P.	
PR	04-JUN-1999;	99US-0137502P.	
PR	07-JUN-1999;	99US-0137724P.	
PR	08-JUN-1999;	99US-0138094P.	
PR	10-JUN-1999;	99US-0138540P.	
PR	10-JUN-1999;	99US-0138847P.	
PR	14-JUN-1999;	99US-0139119P.	
PR	16-JUN-1999;	99US-0139452P.	
PR	16-JUN-1999;	99US-0139453P.	
PR	17-JUN-1999;	99US-0139492P.	
PR	18-JUN-1999;	99US-0139454P.	

PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 28-JUN-1999; 99US-0140655P.
PR 29-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0140991P.
PR 01-JUL-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 12-JUL-1999; 99US-0142920P.
PR 13-JUL-1999; 99US-0142977P.
PR 14-JUL-1999; 99US-0143542P.
PR 15-JUL-1999; 99US-0143624P.
PR 16-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145193P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 27-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147322P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 10-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.

PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0149930P.
PR 26-AUG-1999; 99US-0150566P.
PR 27-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 30-AUG-1999; 99US-0151066P.
PR 31-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151303P.
PR 07-SEP-1999; 99US-0151930P.
PR 10-SEP-1999; 99US-0152363P.
PR 13-SEP-1999; 99US-0153070P.
PR 16-SEP-1999; 99US-0153758P.
PR 16-SEP-1999; 99US-0154018P.
PR 20-SEP-1999; 99US-0154039P.
PR 22-SEP-1999; 99US-0154779P.
PR 23-SEP-1999; 99US-0155139P.
PR 24-SEP-1999; 99US-0155486P.
PR 28-SEP-1999; 99US-0156559P.
PR 29-SEP-1999; 99US-0156458P.
PR 04-OCT-1999; 99US-0156596P.
PR 05-OCT-1999; 99US-0157117P.
PR 06-OCT-1999; 99US-0157753P.
PR 07-OCT-1999; 99US-0157865P.
PR 08-OCT-1999; 99US-0158009P.
PR 12-OCT-1999; 99US-0158232P.
PR 13-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159638P.
PR 21-OCT-1999; 99US-0159844P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 23-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 28-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 29-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 70.2%; Score 59; DB 3; Length 120;
Best Local Similarity 66.7%; Pred. No. 0.25;

Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 HNHNNNNNNH 12
Db 92 NNHSHNNHSHN 103

RESULT 71
ABG12879
ID ABG12879 standard; protein; 124 AA.


```
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 09-JUL-1999; 99US-0142390P.
PR 12-JUL-1999; 99US-0142920P.
PR 13-JUL-1999; 99US-0142977P.
PR 14-JUL-1999; 99US-0143542P.
PR 15-JUL-1999; 99US-0143624P.
PR 16-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144335P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144844P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147320P.
PR 06-AUG-1999; 99US-0147203P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148656P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149368P.
PR 18-AUG-1999; 99US-0149175P.
PR 20-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.

PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 13-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0156589P.
PR 26-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 05-OCT-1999; 99US-0157117P.
PR 06-OCT-1999; 99US-0157753P.
PR 07-OCT-1999; 99US-0157865P.
PR 12-OCT-1999; 99US-0158029P.
PR 13-OCT-1999; 99US-0158232P.
PR 13-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160880P.
PR 22-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0161040P.
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 28-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 70.2%; Score 59; DB 3; Length 135;
Best Local Similarity 66.7%; Pred. No. 0.28;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
DB 107 NNNSHNNHSHN 118

RESULT 73
ABG93228
ID ABG93228 standard; protein: 171 AA.
XX ABG93228;
AC
XX
XX
DT 21-NOV-2002 (first entry)
XX
XX C. albicans BAX-associated protein fragment SEQ ID 414.
DE Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
XX Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
KW Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
KW Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
```

KW	apoptosis; fungal; yeast; infection; autoimmune disease; ischemia;
KW	neurodegeneration; cell death.
XX	
OS	Candida albicans.
XX	
PN	WO200264766-A2.
XX	
PD	22-AUG-2002.
XX	
XX	21-DEC-2001; 2001WO-EP015398.
XX	
PR	22-DEC-2000; 2000EP-00870318.
XX	
PR	04-JAN-2001; 2001EP-00870002.
XX	
PR	09-JAN-2001; 2001EP-00870003.
XX	
XX	
PA	(JANC) JANSSEN PHARM NV.
XX	
PI	Conterras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
XX	
DR	WPI: 2002-667002/71.
XX	
DR	N-PSDB; ABQ76494.
XX	
PT	New isolated nucleic acid representing a synthetic BAX-gene, useful as
XX	medicament for treating, preventing and/or alleviating yeast or fungal
PT	infections or proliferative disorders, or for preventing apoptosis in
XX	certain diseases.
XX	
XX	Claim 36; Fig 2; 344pp; English.
PS	
XX	
CC	This invention describes a novel nucleic acid representing a synthetic
CC	Bax gene. The Bax gene of the invention is useful for identifying Bax-
CC	resistant yeast or fungi, identifying, or obtaining and identifying
CC	Candida spp. sequences that are differentially expressed in a pathway
CC	eventually leading to programmed cell death or identifying inhibitors or
CC	inhibitor sequences of Bax-induced cell death. The products of the
CC	invention have cytosstatic, fungicide; immunosuppressive, virucide and
CC	vasotropic activity and can be used in vaccines or for gene therapy. The
CC	isolated nucleic acids, polypeptides, pharmaceutical compositions,
CC	antisense molecules and antibodies are useful as medicaments or in
CC	preparing a medicament for treating, preventing and/or alleviating
CC	diseases associated with yeast or fungi or proliferative disorders, such
CC	as cancer, or for preventing apoptosis in certain diseases. The compounds
CC	or polypeptides, or the genetically modified organism are useful for
CC	preparing a medicament for modifying the endogenic flora of humans and
CC	other mammals. The vaccine is useful for immunising against yeast or
CC	fungal infections. Apoptosis-related diseases include autoimmune disease,
CC	ischaemia, diseases related with viral infections or neurodegenerations.
CC	This sequence represents a polypeptide associated with the Bax gene
CC	described in the disclosure of the invention
XX	
XX	
SQ	Sequence 171 AA;
Query Match	70.2%; Score 59; DB 5; Length 171;
Best Local Similarity	58.3%; Pred. No. 0.36;
Matches 7; Conservative	5; Mismatches 0; Indels 0; Gaps 0
Oy	1 HHNNNNHHNNHN 12
	: : : : : :
Dd	100 HDHHHHHHHHN 111
RESULT 74	
ID	ABO60371
XX	ABO60371 standard; protein; 186 AA.
XX	
AC	ABO60371;
XX	
DT	29-JUL-2004 (first entry)
XX	
DE	Human genome derived single exon protein #6605.
XX	
XX	Human; gene expression; single exon probe; microarray;
KW	alternative splicing event; genomic alteration.
KW	

XX	Homo sapiens.
OS	US2003194704-A1.
PN	16-OCT-2003.
XX	
PD	03-APR-2002; 2002US-00029386.
XX	
PF	03-APR-2002; 2002US-00029386.
XX	
PR	03-APR-2002; 2002US-00029386.
XX	(PENN/) PENN S G.
PA	(RANK/) RANK D R.
PA	(HANZ/) HANZEL D K.
XX	
PI	Penn SG, Rank DR, Hanzel DK;
DR	WPI, 2004-119264/12.
XX	
PT	New human genome-derived single exon nucleic acid probes useful for human
PT	gene expression analysis, for identifying or characterizing alternative
PT	splicing events, for assessing genomic alterations or as tools for
PT	surveying tissues.
XX	
PS	Claim 45; SEQ ID NO 34005; 80pp; English.
XX	
CC	The invention relates to a nucleic acid probe for measuring human gene
CC	expression, comprising any of the 27,400 fully defined nucleotide
CC	sequences in the specification, or their complements or fragments, and
CC	encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC	fully defined in the specification. The probe is a single exon probe that
CC	hybridises under high stringency conditions to a nucleic acid molecule
CC	expressed in human cells or tissues. Also included are a spatially-
CC	addressable set of single exon nucleic acid probes for measuring human
CC	gene expression (comprising a plurality of single exon nucleic acid
CC	probes cited above, where each of the plurality of probes is separately
CC	and addressably isolatable or amplifiable from the plurality), a single
CC	exon microarray for measuring human gene expression, a method of
CC	measuring human gene expression, a vector comprising the single exon
CC	probe cited above, an ORF-encoded peptide comprising at least 8
CC	contiguous amino acids of any of the above-mentioned amino acid
CC	sequences (optionally with conservative amino acid substitutions), an
CC	isolated antibody that binds specifically to a peptide cited above,
CC	methods of selling and/or licensing single exon probes or microarrays to
CC	a customer desiring to measure gene expression, a method of providing
CC	human gene expression data by subscription, and a computer-readable
CC	storage medium which contains a database having a plurality of records
CC	(each record including data on the expression of a single exon probe
CC	cited above. The probe, methods and apparatus are useful in gene
CC	expression analysis. The probes may be used as tools for surveying
CC	tissues to detect the presence of expressed messages that contain their
CC	specific exon, or in constructing genome-derived single exon microarrays.
CC	In addition, the probes are used in identifying and characterising
CC	alternative splicing events, in detecting and characterising gross
CC	alterations in the genomic locus that includes their exon, in assessing
CC	smaller genomic alterations, in priming the synthesis of nucleic acids,
CC	or in expressing the ORF-encoded peptide. The present sequence data for
CC	this patent did not form part of the printed specification, but was
CC	obtained in electronic format directly from USFTO at
CC	seqdata.uspto.gov/sequence.html?docid=20030194704
XX	
SQ	Sequence 186 AA:
Query Match	70.2%; Score 59; DB 8; Length 186;
Beet Local Similarity	58.3%; Pred. No. 0.4;
Matches	7; Conservative 5; Mismatches 0; Indels 0; Gaps 0.
QY	1 HNNHHNNHNNHN 12 ::: :: :: : db 151 HNNHHNNHNNHN 162

RESULT 75
 ABUS5543
 ID ABUS5543 standard; protein; 281 AA.
 AC ABUS5543;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #21070.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Mycoplasma genitalium.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA39413.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 63467; 176bp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a gene in an operon required for the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining in which the gene to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 281 AA;

Query Match 70.2%; Score 59; DB 6; Length 281;
 Best Local Similarity 72.7%; Pred. No. 0.62;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
 DB 146 HDNNHNNHNNH 156

RESULT 76
 AAM18326
 ID AAM18326 standard; protein; 292 AA.
 AC AAM18326;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Peptide #4760 encoded by probe for measuring cervical gene expression.
 XX
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000670.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-0063236P.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488901/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human cervical epithelial cells.
 XX
 PS Claim 27; SEQ ID NO 23152; 487bp; English.

The present invention relates to human single exon nucleic acid probes (SENPs; see AAM10068-AAM18459). The present sequence is a peptide encoded by one such probe. The SENPs are derived from human HeLa cells. The SENPs can be used to produce a single exon microarray, which can be used for measuring human gene expression in a sample derived from human cervical epithelial cells. By measuring gene expression, the probes are therefore useful in grading and/or staging of diseases of the cervix, notably cervical cancer. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 292 AA;

Query Match 70.2%; Score 59; DB 4; Length 292;
 Best Local Similarity 58.3%; Pred. No. 0.65;
 Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
 DB 56 HNNHNNHNNH 67

RESULT 77

ABB37360
ID ABB37360 standard; peptide; 292 AA.
XX
AC ABB37360;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #4866 encoded by human foetal liver single exon probe.
XX
KM Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX
PT gene expression in human foetal liver.
XX
PS Claim 27; SEQ ID NO 29995; 639pp + Sequence Listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring
XX human gene expression in a sample derived from human foetal liver. The
XX single exon nucleic acid probes may be used for predicting, measuring and
XX displaying gene expression in samples derived from human foetal liver. The
XX present sequence is a peptide encoded by a single exon nucleic acid probe
XX of the invention. Note: The sequence data for this patent did not form
XX part of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pcr_sequences
XX
SQ Sequence 292 AA;
XX
Query Match 70.2%; Score 59; DB 4; Length 292;
Best Local Similarity 58.3%; Pred. No. 0.65;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
OY 1 HHHNNHHNNHN 12
|:|:|:|:|:|:
Db 56 HHHNNHHHHHH 67
XX
RESULT 78
AAM30814
ID AAM30814 standard; protein; 292 AA.
XX
AC AAM30814;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #4851 encoded by probe for measuring placental gene expression.
XX
KM Probe; microarray; human; placenta; antenatal diagnosis;
XX
KM genetic disorder.
XX
OS Homo sapiens.
XX

PN WO200157272-A2.
XX
XX 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX
PT gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 31083; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
XX see AAI31315-AA157546). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders
XX
SQ Sequence 292 AA;
XX
Query Match 70.2%; Score 59; DB 4; Length 292;
Best Local Similarity 58.3%; Pred. No. 0.65;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
OY 1 HHHNNHHNNHN 12
|:|:|:|:|:|:
Db 56 HHHNNHHHHHH 67
XX
RESULT 79
ABB32108
ID ABB32108 standard; peptide; 292 AA.
XX
AC ABB32108;
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #4759 encoded by breast cell single exon nucleic acid probe.
XX
KM Human; microarray; single exon probe; gene expression; breast; disease;
XX
KM cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-496933/54.
XX PT New spatially-addressable set of single exon nucleic acid probes, useful
XX PT for measuring gene expression in sample derived from human breast,
XX PT comprises number of single exon nucleic acid probes.
XX PS Claim 27; SEQ ID NO 15076; 327bp + Sequence listing; English.
XX XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC assessing breast disease. Gene expression analysis is useful for
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a peptide
CC encoded by a single exon nucleic acid probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX XX
SQ Sequence 292 AA;
Query Match 70.2%; Score 59; DB 4; Length 292;
Best Local Similarity 58.3%; Pred. No. 0.65;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHHNNHHNNHN 12
Db 56 HHHNNHHNNHH 67
RESULT 80
ABB22646
ID ABB22646 standard; protein; 292 AA.
XX AC ABB22646;
XX XX
DT 23-JAN-2002 (first entry)
XX DE Protein #4645 encoded by probe for measuring heart cell gene expression.
XX KW Human; gene expression; heart; microarray; vascular system;
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease.
XX OS Homo sapiens.
XX OS
PN WO200157274-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000666.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488899/53.
XX XX
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT hearts.
XX PS Claim 15; SEQ ID NO 24416; 530bp; English.
XX XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC AB21535-AB241105). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX XX
SQ Sequence 292 AA;
Query Match 70.2%; Score 59; DB 4; Length 292;
Best Local Similarity 58.3%; Pred. No. 0.65;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHHNNHHNNHN 12
Db 56 HHHNNHHNNHH 67
RESULT 81
AAM70490
ID AAM70490 standard; protein; 292 AA.
XX AC AAM70490;
XX XX
DT 06-NOV-2001 (first entry)
XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 30796.
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX OS Homo sapiens.
XX OS
PN WO200157276-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000668.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488900/53.
XX XX
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human bone marrow.
XX PS Example 4; SEQ ID NO 30796; 658bp + Sequence listing; English.
XX XX

CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention

XX Sequence 292 AA;

Query Match 70.2%; Score 59; DB 4; Length 292;
 Best Local Similarity 58.3%; Pred. No. 0.65;
 Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
 |||||:|:|:|:
 Db 56 HHHHHHHHHHHH 67

RESULT 82

AAMS8050
 ID AAMS8050 standard; protein; 292 AA.

AC AAMS8050;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30155.

XX Human; brain expressed exon; gene expression analysis; probe; microarray;

KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

OS Homo sapiens.

PN MO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000667.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLF-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT brains.

PS Example 4; SEQ ID NO 30155; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is a protein encoded by one of

CC the probes of the invention

XX Sequence 292 AA;

QY Query Match 70.2%; Score 59; DB 4; Length 292;
 Best Local Similarity 58.3%; Pred. No. 0.65;
 Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12

Db 56 HHHHHHHHHHHH 67
 |||||:|:|:|:

RESULT 83

ABG52171
 ID ABG52171 standard; peptide; 292 AA.

AC ABG52171;

DT 25-FEB-2003 (first entry)

DE Human liver peptide, SEQ ID NO 30819.

XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;

KW hypercholesterolaemia; coronary heart disease.

OS Homo sapiens.

PN MO200157273-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000664.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLF-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-48898/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human adult liver.

PS Claim 27; SEQ ID NO 30819; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for

CC measuring human gene expression in a sample derived from human adult

CC liver, comprising one of 13109 defined nucleotide sequences given in the

CC specification (or complements/ fragments). The probe hybridises at high

CC stringency to a nucleic acid molecule expressed in the human adult liver.

CC (I) may be used for predicting, measuring and displaying gene expression

CC in samples derived from human adult liver. The genes identified may be

CC involved in genetic liver diseases such as cirrhosis,

CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is

CC associated with coronary heart disease. ABG47348-ABG5930 represent human

CC liver single exon encoded peptides of the invention. Note: The sequence

CC information for this patent does not appear in the printed specification

CC but was obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 292 AA;

QY Query Match 70.2%; Score 59; DB 4; Length 292;
 Best Local Similarity 58.3%; Pred. No. 0.65;
 Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
 |||||:|:|:|:
 Db 56 HHHHHHHHHHHH 67

RESULT 84
 AAM05933
 ID AAM05933 standard; protein; 292 AA.

XX
AC AAM05933;
XX
DT 09-OCT-2001 (first entry)
XX
DE Peptide #4615 encoded by probe for measuring breast gene expression.
XX
KW Probe; human; breast disease; breast cancer; development disorder;
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US000661.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression in
XX a human breast.
XX
PS Claim 27; SEQ ID NO 14673; 322pp; English.
XX
SQ The present invention relates to novel single exon nucleic acid probes
CC (see AAT00010-AAT10067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 292 AA;
XX
Query March 70.2%; Score 59; DB 4; Length 292;
Best Local Similarity 58.3%; Pred. No. 0.65;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 HHNNHHNNHHNN 12
Db |||||:||||:
56 HHNNHHHHHHHH 67
XX
RESULT 85
ABG40129
XX ABG40129 standard; peptide; 292 AA.
XX
AC ABG40129;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 29794.
XX
KW Human; single exon probe; asthma; lung cancer; COPD; IID;

KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsagen syndrome;
KW primary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX
PS Claim 27; SEQ ID NO 29794; 634pp; English.
XX
SQ The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridises at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC collected from human lung, comprising (a) contacting the array with a
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryotic lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarray having a probe in the tissues and/or cell types indicates that
CC expression of the exons in the tissues and/or cell types indicates that
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karsagen syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a peptide/protein encoded by a single exon probe of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 292 AA;

Query Match 70.2%; Score 59; DB 5; Length 292;
 Best Local Similarity 58.3%; Pred. No. 0.65;
 Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
 |:|:|:|:|:|:
 Db 56 HHHHHHHHHH 67

RESULT 86

ABB61270
 ID ABB61270 standard; protein; 299 AA.

AC ABB61270;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 10602.

KW Drosophila; developmental biology; cell signalling; insecticide;
 pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.
 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL05373.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.

PS Disclosure; SEQ ID NO 10602; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL1840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 299 AA;

Query Match 70.2%; Score 59; DB 4; Length 299;
 Best Local Similarity 58.3%; Pred. No. 0.66;

Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
 |:|:|:|:|:|:
 Db 115 HHHHHHHHHH 126

RESULT 87

ABB60569
 ID ABB60569 standard; protein; 374 AA.

AC ABB60569;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 8499.

KW Drosophila; developmental biology; cell signalling; insecticide;
 pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.
 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL04672.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.

PS Disclosure; SEQ ID NO 8499; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL1840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 374 AA;

Query Match 70.2%; Score 59; DB 4; Length 374;
 Best Local Similarity 58.3%; Pred. No. 0.84;
 Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
 |:|:|:|:|:|:
 Db 164 HSHHSHDHHDH 175

RESULT 88

ADP99129
 ID ADP99129 standard; protein; 467 AA.

AC ADP99129;

DT 23-SEP-2004 (first entry)

DE C. albicans specific gene, orf19.5065, protein sequence.

XX Diploid fungal cell; allele; gene disruption cassette;
 KW promoter replacement fragment; antifungal; fungicide; gene therapy;
 KW infection; Candida albicans.

XX Candida albicans.

WO2004056965-A2.
 08-JUL-2004.
 19-DEC-2003; 2003WO-US040618.
 19-DEC-2002; 2002US-0434832P.
 (ELIT-) ELITRA PHARM INC.
 (ELIT-) ELITRA CANADA LTD.
 Roemer T, Jiang B, Boone C, Bussey H;
 WPI: 2004-500296/47.
 N-PSDB; ADP98819.
 Constructing a strain of diploid fungal cells in which both alleles of a gene are modified comprises modifying the alleles of a gene in the fungal cells by recombination using a gene disruption cassette and a promoter replacement fragment.
 Claim 44; SEQ ID NO 7304; 163bp; English.
 The invention relates to a novel method for constructing a strain of diploid fungal cells in which both alleles of a gene are modified. The method comprises modifying the alleles of a gene in diploid fungal cells by recombination using a gene disruption cassette and a promoter replacement fragment. The invention further comprises: assembling a collection of diploid fungal cells each of which comprises modified alleles of a different gene; a strain of diploid fungal cells comprising inactivated by a gene disruption cassette comprising a nucleotide sequence encoding an expressible selectable marker; and the expression of that is operably linked to the coding region of the second allele of the gene, and where the gene encodes the polypeptide mentioned above; a collection of diploid fungal strains comprising the diploid strains cited above, where substantially all the different genes that encode the above amino acid sequences are modified and are present in different diploid strains in the collection; a nucleic acid molecule microarray comprising nucleic acid molecules, where each nucleic acid molecule comprises a nucleotide sequence that is hybridizable to a target nucleotide sequence comprising any of the 310 nucleotide sequences listed in the specification (ADP98516-ADP98825); identifying a gene that is essential to the survival or growth of a fungus, that contributes to the virulence and/or pathogenicity of a fungus, or that contributes to the resistance of a diploid fungus to an antifungal agent; identifying an antifungal agent that inhibits the growth of a diploid fungus; or a therapeutic level of treatment of a mammalian disease; correlating changes in the proliferation of a diploid fungal cell; a purified or isolated nucleic acid molecule comprising a nucleotide sequence encoding a gene product required for proliferation of *Candida albicans*, where the gene product consists of any of the above-mentioned amino acid sequences; a vector comprising a promoter operably linked to the nucleic acid molecule cited above; a host cell containing the vector; a purified or isolated polypeptide comprising any of the 61 amino acid sequences given in the specification (ADP96718-ADP96778); a fusion protein comprising a fragment of a first polypeptide fused to a second polypeptide, the fragment consisting of at least 6 consecutive residues of any of ADP98826-ADP9915 activity of a gene product; identifying a compound which modulates the activity of a gene product encoded by a nucleic acid comprising any of ADP98516-ADP98825, eliciting an immune response in an animal; a strain of *Candida albicans*, where a first allele of a gene comprising any of ADP98516-ADP98825 is inactive and a second allele of the gene is under the control of a heterologous promoter; identifying a compound or binding partner that binds to the polypeptide comprising any of ADP98826-ADP9915, or its fragment; identifying a compound having the ability to inhibit growth or proliferation of *Candida albicans*; inhibiting growth or proliferation of *Candida albicans* cells; manufacturing an antimycotic compound; treating an infection of a subject by *Candida albicans*; preventing or containing contamination of an object by *Candida albicans*;

or for preventing or inhibiting formation on a surface of a biofilm comprising *Candida albicans*; a pharmaceutical composition comprising a therapeutic amount of an agent which reduces the activity or level of a gene product encoded by a nucleic acid comprising any of ADP98516-ADP98825 in a pharmaceutical carrier; an antibody preparation which binds the polypeptide; methods for evaluating a compound against a target gene product encoded by any of ADP98516-ADP98825; identifying an antimycotic compound; a computer or a computer readable medium that comprises at least one of the nucleotide sequences mentioned in the specification or method assisted by a computer for identifying a putatively essential gene of a fungus; and a protein array comprising proteins, where at least one sequence selected from ADP98516-ADP98825, where at least one composition have fungicide activity. The novel methods and gene therapy. The composition and methods are useful for drug screening purposes or for diagnosing, preventing or treating infections associated with *Candida albicans*. These may also be used for constructing strains useful for identification and validation of gene products as effective targets for therapeutic intervention, for identifying and validating gene products as effective targets for therapeutic intervention, and for collecting identified essential genes. This sequence represents the protein of a *Candida albicans* fungal specific gene of the invention. NOTE: This sequence was downloaded from an electronic sequence listing provided on the WIPO website.

Sequence 467 AA;

Query Match 70.2%; Score 59; DB 8; Length 467;
 Best Local Similarity 81.8%; Prod. No 1.1;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNNNNNNNN 12
 Db 78 NNNNNNNNN 88

RESULT 89
 ABB59185
 ID ABB59185 standard; protein; 537 AA.
 AC ABB59185;
 XX

26-MAR-2002 (first entry)

Drosophila melanogaster polypeptide SEQ ID NO 4347.

Drosophila; developmental biology; cell signalling; insecticide;
 pharmaceutical.

Drosophila melanogaster.

WO200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001WO-US009231.

23-MAR-2000; 2000US-0191637P.

11-JUL-2000; 2000US-00614150.

(PERK) PE CORP NY.

Venter JC, Adams M, Li PWD, Myers EW;

WPI: 2001-656860/75.
 N-PSDB; ABL03288.

New isolated nucleic acid detection reagent for detecting 1000 or more genes from *Drosophila* and for elucidating cell signalling and cell-cell interactions.

Discloure; SEQ ID NO 4347; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 537 AA;

Query Match 70.2%; Score 59; DB 4; Length 537;
 Best Local Similarity 58.3%; Pred. No. 1.2;
 Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
 |::|::|::|::|
 Db 355 HSHSHSHHHHNN 366

RESULT 90
 ABB64623
 ID ABB64623 standard; protein; 989 AA.

XX ABB64623;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 20661.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PMD, Myers EW;

XX WPI, 2001-656860/75.

XX N-PSDB; ABL08726.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.

XX Disclosure; SEQ ID NO 20661; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 989 AA;

Query Match 70.2%; Score 59; DB 4; Length 989;
 Best Local Similarity 58.3%; Pred. No. 2.4;
 Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
 |::|::|::|::|
 Db 188 HHHHHHHHHHNN 199

RESULT 91
 ABB68075
 ID ABB68075 standard; protein; 1300 AA.

XX ABB68075;
 XX

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 31017.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PMD, Myers EW;

XX WPI, 2001-656860/75.

XX N-PSDB; ABL12178.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.

XX Disclosure; SEQ ID NO 31017; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1300 AA;

Query Match 70.2%; Score 59; DB 4; Length 1300;
 Best Local Similarity 58.3%; Pred. No. 3.2;
 Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
 |::|::|::|::|
 Db 544 HHHHHHHHHHNN 555

RESULT 92
 AAW58576
 ID AAW58576 standard; peptide; 19 AA.
 XX
 AC AAW58576;

```

XX 17-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 12-OCT-1998 (first entry)
XX
DE ACEII His-rich peptide.
XX
KW ACEI; transcriptional activator protein; acel gene; cellulase;
KW cellobiohydrolase; binuclear cluster domain.
XX
OS Hypocrea jecorina.
XX
PN WO9823642-A1.
XX
PD 04-JUN-1998.
XX
PF 01-DEC-1997; 97MO-FI000743.
XX
PR 29-NOV-1996; 96US-0032156P.
PR 13-DEC-1996; 96US-0032259P.
PR 10-MAR-1997; 97US-0040140P.
XX
PA (ROHG ) ROEHM ENZYME FINLAND OY.
XX
PI Saloheimo A, Aro N, Ilmen M, Penttilae M;
DR WPI; 1998-322661/28.
XX
PT New isolated transcriptional regulatory protein nucleic acid(s) - used
PT to, e.g. enhance production of homologous and heterologous proteins in
PT fungal hosts.
XX
PS Example 5; Page 32; 99p; English.
XX
CC This peptide corresponds to amino acid residues 47-65 of ACEII (see
CC AAM8573), a novel transcriptional regulatory protein of Trichoderma
CC reesei that is capable of activating the cellulase cbh1 gene promoter.
CC This region of ACEII follows the DNA binding domain and shows similarity
CC to some homeobox-containing regulatory proteins of Drosophila
CC melanogaster. Methods are provided for using ACEII sequences to enhance
CC production of homologous or heterologous proteins, especially enzymes, in
CC fungal hosts. (Updated on 25-MAR-2003 to correct PA field.) (Updated on
CC 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 19 AA;
Query Match 69.0%; Score 58; DB 2; Length 19;
Best Local Similarity 58.3%; Pred. No. 0.047;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 HNNHNNHNNHNN 12
Db 7 HSHSHSHSHSHN 18
RESULT 93
ID ABG10513
XX ABG10513 standard; protein; 51 AA.
XX
AC ABG10513;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #10504.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.

```

```

XX 30-MAR-2001; 2001WO-US008631.
PF
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX
DR N-PSDB; AAS74700.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostic, forensic, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 40872; 103p; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (II) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensic, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC patent did not appear in the invention. Note: The sequence data for this
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 51 AA;
Query Match 69.0%; Score 58; DB 4; Length 51;
Best Local Similarity 63.6%; Pred. No. 0.14;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 HNNHNNHNNHNN 11
Db 7 HHHHHHHNNHH 17
RESULT 94
ID ABG26844
XX ABG26844 standard; protein; 117 AA.
XX
AC ABG26844;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #26835.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
PF 30-MAR-2001; 2001WO-US008631.

```

XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX N-PSDB; AAS91031.
PT
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 57203; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridization probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 117 AA;
Query Match 69.0%; Score 58; DB 4; Length 117;
Best Local Similarity 58.3%; Pred. No. 0.33;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
QY 1 HNNNNNNNNNN 12
:|:|:|:|:|:|:
Db 59 YNNSHHHHHHH 70
RESULT 95
AAM21094
ID AAM21094 standard; protein; 156 AA.
XX
AC AAM21094;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #7558 encoded by probe for measuring cervical gene expression.
XX
XX Probe; human; microarray; gene expression; cervical epithelial cell;
KM cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX

PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
PT
PS Claim 27; SEQ ID NO 25920; 487bp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SENPs: see A110068-A128459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 156 AA;
Query Match 69.0%; Score 58; DB 4; Length 156;
Best Local Similarity 58.3%; Pred. No. 0.45;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 HNNNNNNNNNN 12
:|:|:|:|:|:|:
Db 100 HHHHHHHHHQH 111
RESULT 96
ABB43409
ID ABB43409 standard; peptide; 156 AA.
XX
AC ABB43409;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #10915 encoded by human foetal liver single exon probe.
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
KM
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX

XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 27; SEQ ID NO 36044; 639bp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human fetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 156 AA;
XX
Query Match 69.0%; Score 58; DB 4; Length 156;
Best Local Similarity 58.3%; Pred. No. 0.45;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
XX
OY 1 HNNNNNNNNNN 12
Db 100 HNNNNNNNNNN 111
XX
RESULT 97
AAM37293
ID AAM37293 standard; protein; 156 AA.
XX
AC AAM37293;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #11330 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488997/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 37562; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP;
CC see AAT1315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX

SQ Sequence 156 AA;
XX
Query Match 69.0%; Score 58; DB 4; Length 156;
Best Local Similarity 58.3%; Pred. No. 0.45;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
XX
OY 1 HNNNNNNNNNN 12
Db 100 HNNNNNNNNNN 111
XX
RESULT 98
ABB26381
ID ABB26381 standard; protein; 156 AA.
XX
AC ABB26381;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #8380 encoded by probe for measuring heart cell gene expression.
XX
DE Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488999/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 15; SEQ ID NO 28151; 530bp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 156 AA;
XX
Query Match 69.0%; Score 58; DB 4; Length 156;
Best Local Similarity 58.3%; Pred. No. 0.45;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
XX
OY 1 HNNNNNNNNNN 12
Db 100 HNNNNNNNNNN 111
XX

Db 100 HHHHHHHHHQH 111

RESULT 99

ABG46175

ID ABG46175 standard; peptide; 156 AA.

XX ABG46175;

AC

XX

DT 19-AUG-2002 (first entry)

XX

DE Human peptide encoded by genome-derived single exon probe SEQ ID 35840.

XX

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;

KM chronic obstructive pulmonary disease; interstitial lung disease;

KM familial idiopathic pulmonary fibrosis; neurofibromatosis;

KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KM Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

KM pulmonary histiocytosis; lymphangioleiomyomatosis; Kartagener syndrome;

KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KM primary ciliary dyskinesia; pulmonary hypertension;

KM hyaline membrane disease.

XX

OS Homo sapiens.

XX

PN MO200186003-A2.

XX

PD 15-NOV-2001.

XX

PF 30-JAN-2001; 2001WO-US000665.

XX

PR 04-FEB-2000; 2000US-0180312P.

XX

PR 26-MAY-2000; 2000US-0207456P.

XX

PR 30-JUN-2000; 2000US-00608408.

XX

PR 03-AUG-2000; 2000US-00632366.

XX

PR 21-SEP-2000; 2000US-0234687P.

XX

PR 27-SEP-2000; 2000US-0236359P.

XX

PR 04-OCT-2000; 2000GB-00024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2002-114183/15.

XX

PT Spatially-addressable set of single exon nucleic acid probes, used to

XX

PS measure gene expression in human lung samples.

XX

XX Claim 27; SEQ ID NO 35840; 634bp; English.

XX

XX The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived

CC from human lung comprising single exon nucleic acid probes having one of

CC 12614 nucleic acid sequences mentioned in the specification, or their

CC complements or the 12387 open reading frames derived from the 12614

CC probes. Also included are a microarray comprising the novel set of probes

CC; the novel set of probes which hybridise at high stringency to a nucleic

CC acid expressed in the human lung; measuring gene expression in a sample

CC derived from human lung; comprising (a) contacting the array with a

CC collection of detectably labeled nucleic acids derived from human lung

CC mRNA, and (b) measuring the label detectably bound to each probe of the

CC array; identifying exons in a eukaryotic genome, comprising (a)

CC algorithmically predicting at least one exon from genomic sequences of

CC the eukaryote; and (b) detecting specific hybridisation of detectably

CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

CC having a fragment identical to the predicted exon, the probe is included

CC in the above mentioned microarray; assigning exons to a single gene,

CC comprising (a) identifying exons from genomic sequence by the method

CC above and (b) measuring the expression of each of the exons in several

CC tissues and/or cell types using hybridisation to a single exon

CC microarrays having a probe with the exon, where a common pattern of

CC expression of the exons in the tissues and/or cell types indicates that

CC the exons should be assigned to a single gene; a peptide comprising one

CC of 12011 sequences, mentioned in the specification, or encoded by the

CC probes/open reading frames (ORF). The probes are used for gene expression

CC analysis, and for identifying exons in a gene, particularly using human

CC lung derived mRNA and for the study of lung diseases such as asthma, lung

CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung

CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,

CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-

CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary

CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,

CC Kartagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary

CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The

CC present sequence is a peptide/protein encoded by a single exon probe of

CC the invention. Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SO Sequence 156 AA;

XX

QY

1 HHHHHHHHHHHN 12

Db 100 HHHHHHHHHQH 111

Query Match 69.0%; Score 58; DB 5; Length 156;

Best Local Similarity 58.3%; Pred. No. 0.45;

Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

XX

RESULT 100

ADG38684

ID ADG38684 standard; protein; 164 AA.

XX

AC ADG38684;

XX

DT 18-DEC-2003 (first entry)

XX

DE Human secreted protein #21.

XX

KM

KM immune disorder; severe combined immunodeficiency; SCID;

KM autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;

KM rheumatoid arthritis; allergic reaction; asthma; myeloid cell deficiency;

KM lymphoid cell deficiency; osteoporosis; osteoarthritis;

KM peripheral nervous system disease; peripheral neuropathy;

KM Alzheimer's disease; Parkinson's disease; coagulation disorder;

KM inflammatory disease; systemic inflammatory response syndrome; SIRS;

KM ischaemia-reperfusion injury; Crohn's disease; anaphylaxis;

KM hypersensitivity; regeneration; neural cell proliferation; fertility;

KM tumour; chemokine; human; secreted protein.

XX

XX Homo sapiens.

OS

PN US2002193567-A1.

XX

PD 19-DEC-2002.

XX

PF 02-APR-2002; 2002US-00114893.

XX

PR 11-AUG-1995; 95US-00514014.

XX

PR 05-APR-1996; 96US-00628364.

XX

PR 19-APR-1996; 96US-00635311.

XX

PR 07-JUN-1996; 96US-00659224.

XX

PR 17-JUN-1996; 96US-00664596.

XX

PR 09-JUL-1996; 96US-00677231.

XX

PR 26-JUL-1996; 96US-00686878.

XX

PR 23-AUG-1996; 96US-00701819.

XX

PR 27-SEP-1996; 96US-00721488.

XX

PR 27-SEP-1996; 96US-00721798.

XX

PR 27-SEP-1996; 96US-00721923.

XX

PR 27-SEP-1996; 96US-00721926.

XX

PR 27-SEP-1996; 96US-00738367.

XX

PR 30-OCT-1996; 96US-00738375.

XX

PR 13-JAN-1997; 97US-00783395.

XX

PR 10-APR-1997; 97US-00833823.

XX

PR 02-JUN-1997; 97US-00867677.

PR 05-SEP-1997; 97US-00924838.
 PR 06-OCT-1999; 99US-00413232.
 XX
 PA (GEWY) GENETICS INST INC.
 XX

PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Bowman MR, Spaulding V, Carlin-Duckett M;
 PI Kellner K;
 XX

DR WPI; 2003-657236/62.
 DR N-PSDB; ADCC8683.
 XX

PT Proteins AZ3021 encoded by clone AZ3021 from human adult colon, and
 PT BD12716 encoded by clone BD12716 from human fetal kidney cDNA library,
 XX useful for treating e.g. multiple sclerosis and rheumatoid arthritis.
 PS Disclosure; SEQ ID NO 42; 412pp; English.

XX
 CC The invention relates to a protein comprising fully defined AZ302 1
 CC protein or BD127 1 6 protein. The polynucleotides are useful for
 CC expressing recombinant proteins for analysis and are also useful as
 CC chromosome markers or tags to identify chromosomes or to map related gene
 CC positions. The proteins are useful as amino acid supplement, carbon
 CC source, nitrogen source and carbohydrate source. The proteins are useful
 CC for treating various immune deficiencies and disorders (e.g. severe
 CC combined immunodeficiency (SCID), autoimmune disorders (e.g. multiple
 CC sclerosis, systemic lupus erythematosus, rheumatoid arthritis), allergic
 CC reactions (e.g. asthma), myeloid or lymphoid cell deficiencies,
 CC osteoporosis or osteoarthritis, peripheral nervous system diseases (e.g.
 CC peripheral neuropathy, Alzheimer's disease, Parkinson's disease),
 CC coagulation disorders, inflammatory diseases (e.g. systemic inflammatory
 CC response syndrome (SIRS), ischemia-reperfusion injury, Crohn's disease),
 CC anaphylaxis and hypersensitivity. Proteins are also useful for inducing
 CC tumour immunity, for inducing bone, cartilage, tendon, ligament and/or
 CC nerve growth or regeneration, for proliferating neural cells and for
 CC regenerating nerve and brain tissue, for inducing fertility and for
 CC inhibiting tumour growth. Proteins are also useful as chemokine for
 CC mammalian cells (e.g., monocytes, fibroblasts, neutrophils), and also
 CC useful as inhibitors of receptor/ligand interactions. The present
 CC sequence represents the amino acid sequence of a human secreted protein.
 XX

Sequence 164 AA:

Query Match
 Best Local Similarity 69.0%; Score 58; DB 7; Length 164;
 Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
 :|:|:|:|:|:|:
 Db 18 YNHSNNHHHHHH 29

Search completed: October 4, 2005, 11:50:14
 Job time : 129 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM,protein - protein search, using sw model

Run on: October 4, 2005, 11:22:37 ; Search time 48 Seconds
(without alignments)
24.054 Million cell updates/sec

Title: US-09-858-332G-15
Perfect score: 84
Sequence: 1 HHHNNHHNNHHN 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	92.9	859	2 S64195	HTRI protein - yea
2	76	90.5	102	2 T02917	hypothetical prote
3	73.5	87.5	1671	2 S71628	sensory transducti
4	70	83.3	490	2 A46391	CAMP receptor subt
5	67	79.8	898	2 T49492	hypothetical prote
6	65	77.4	303	2 AG1396	cation transport p
7	65	77.4	303	2 AB1772	probable homeodoma
8	64	76.2	244	2 A84687	conserved hypothet
9	64	76.2	451	2 G70176	hypothetical prote
10	64	76.2	529	2 T00677	hypothetical prote
11	62	73.8	329	2 T45972	hypothetical prote
12	62	73.8	356	2 T43145	hypothetical prote
13	62	73.8	529	2 T08684	hypothetical prote
14	62	73.8	601	2 T00119	probable transcrip
15	61	72.6	126	2 T15993	hypothetical prote
16	61	72.6	190	2 D64087	probable peptidylp
17	61	72.6	1028	2 A56038	DNA-binding protei
18	61	72.6	1213	2 S16356	ovo protein - frui
19	59	70.2	180	2 B97242	hypothetical prote
20	59	70.2	200	2 G97048	probable membrane
21	59	70.2	281	2 E64216	hypothetical prote
22	59	70.2	375	2 A46390	CAMP receptor subt
23	59	70.2	394	2 C84905	probable extensin
24	59	70.2	434	2 I51436	HNF-3beta - Africa
25	59	70.2	444	2 F96836	hypothetical prote
26	59	70.2	580	2 T46024	hypothetical prote
27	59	70.2	1291	2 T13389	hypothetical prote
28	58	69.0	719	2 S61046	ARPI protein - yea
29	57	67.9	259	2 A70359	hydrogenase expres

30	57	67.9	285	2 E84766	probable AT-hook D
31	57	67.9	477	2 T47753	hypothetical prote
32	57	67.9	658	2 T04219	hypothetical prote
33	57	67.9	731	2 S46115	hypothetical prote
34	56	66.7	107	2 H86642	hypothetical prote
35	56	66.7	212	2 S49807	hypothetical prote
36	56	66.7	360	2 T19585	hypothetical prote
37	56	66.7	385	2 A84696	probable zinc tran
38	56	66.7	466	2 S19365	hypothetical prote
39	56	66.7	608	2 T18437	hypothetical prote
40	56	66.7	828	2 C88402	protein H05C05.1 f
41	56	66.7	1711	2 T18429	hypothetical prote
42	55	65.5	381	2 A47327	selenoprotein P pr
43	55	65.5	386	2 A96625	hypothetical prote
44	55	65.5	740	2 S61568	probable membrane
45	55	65.5	786	2 T18469	hypothetical prote
46	54	64.3	172	2 T51065	hypothetical prote
47	54	64.3	236	2 S41512	Bm-3b protein - m
48	54	64.3	303	2 A56837	homeotic protein M
49	54	64.3	303	2 B49122	growth arrest-spec
50	54	64.3	303	2 T44684	hypothetical prote
51	54	64.3	306	2 T44684	hypothetical prote
52	54	64.3	308	2 E81923	probable periplasm
53	54	64.3	326	2 D83483	probable metal tra
54	54	64.3	349	2 B95858	conserved hypothet
55	54	64.3	391	2 H86187	hypothetical prote
56	54	64.3	407	2 G84783	probable peptidest
57	54	64.3	410	2 T18502	gene Bm-3b protei
58	54	64.3	411	2 T58156	Bm-3.2 - mouse
59	54	64.3	414	2 A48273	delta/Y11/NF-E1/UC
60	54	64.3	420	2 T39712	hypothetical prote
61	54	64.3	427	2 T42516	hypothetical prote
62	54	64.3	430	2 S66671	neuron-derived rec
63	54	64.3	450	2 C96704	unknown protein, 2
64	54	64.3	515	2 S52453	ecdysteroid UDP-gl
65	54	64.3	530	2 S66148	gene pipsqueak pro
66	54	64.3	535	2 T49672	related to a-aglu
67	54	64.3	628	2 J71930	neuron-derived rec
68	54	64.3	628	2 JC2493	neuron derived orp
69	54	64.3	633	1 A26030	serine/threonine-s
70	54	64.3	754	2 UC4898	Down-syndrome-crit
71	54	64.3	813	2 AH3258	cation-transportin
72	54	64.3	1085	2 S66149	gene pipsqueak pro
73	54	64.3	1180	2 S69205	stripe a/b protein
74	54	64.3	1256	2 S14556	asparagine-rich pr
75	54	64.3	1273	2 T00338	hypothetical prote
76	53	63.1	60	2 C64698	probable histidine
77	53	63.1	133	2 B30242	stem cell protein
78	53	63.1	143	2 B64421	conserved hypothet
79	53	63.1	185	2 G64075	urase accessory p
80	53	63.1	270	2 S74993	hypothetical prote
81	53	63.1	279	2 H90992	hypothetical prote
82	53	63.1	283	2 C85838	hypothetical prote
83	53	63.1	302	2 A55641	homeotic protein G
84	53	63.1	311	2 A56235	transcription acti
85	53	63.1	331	2 A30242	homeotic protein E
86	53	63.1	337	2 D64049	adhesin homolog Hi
87	53	63.1	341	2 E83340	hypothetical prote
88	53	63.1	351	1 KGZ0HL	histidine-rich gly
89	53	63.1	353	2 AB2396	hypothetical prote
90	53	63.1	355	2 S35345	oxl protein - mou
91	53	63.1	355	2 I56547	homeodomain protei
92	53	63.1	383	2 S76964	hypothetical prote
93	53	63.1	414	2 A40350	transcription repr
94	53	63.1	420	2 C96995	uncharacterized co
95	53	63.1	469	2 I37451	HBR-G2 (HBR-2) pro
96	53	63.1	472	2 A54743	transcription fact
97	53	63.1	622	2 S71342	calnexin precursor
98	53	63.1	623	1 S31167	gene pointed prote
99	53	63.1	657	2 A29454	knob-associated hi
100	53	63.1	669	2 S14535	asparagine-rich pr

ALIGNMENTS

RESULT 1

S64195

HTR1 protein - yeast (Saccharomyces cerevisiae)

N.Alternate names: protein BIC834; protein G1625; protein YGL178W; suppressor protein MF

C.Species: Saccharomyces cerevisiae

C.Date: 17-May-1996 #sequence revision 17-May-1996 #text change 09-Jul-2004

C.Accession: S64195; S51561; S50194; S57253; S56166; S50156

R.Brusch, C.V.; Coglievina, M.; Bertani, I.; Klima, R.; Zaccaria, P.; Delneri, D.

submitted to the Protein Sequence Database, May 1996

A.Reference number: S64193

A.Accession: S64195

A.Molecule type: DNA

A.Residues: 1-859 <BRU>

A.Cross-references: UNIPROT:P39016; EMBL:Z72700; NID:g1322785; PIDN:CAA96869.1; PID:e243

A.Experimental source: strain S288C

R.Kikuchi, Y.; Oka, Y.; Kobayashi, M.; Uesono, Y.; Toh-e, A.; Kikuchi, A.

Mol. Gen. Genet. 245, 107-116, 1994

A>Title: A new yeast gene, HTR1, required for growth at high temperature, is needed for

A.Reference number: S51561; MUID:95147841; PMID:7845352

A.Accession: S51561

A.Molecule type: DNA

A.Residues: 26-859 <KIK>

A.Cross-references: EMBL:D25541; NID:g436256; PIDN:BA05024.1; PID:d1005565; PID:g469512

R.Sakai, A.

submitted to the EMBL Data Library, December 1993

A.Description: Multicopy suppressors of the yeast pop2 mutation.

A.Reference number: S48511

A.Accession: S50194

A.Molecule type: DNA

A.Residues: 26-859 <SAX>

A.Cross-references: EMBL:D26184; NID:g450489; PIDN:BA05172.1; PID:d1005714; PID:g450490

R.Coglievina, M.; Bertani, I.; Klima, R.; Zaccaria, P.; Brusch, C.V.

Yeast 11, 767-774, 1995

A>Title: The DNA sequence of a 7941 bp fragment of the left arm of chromosome VII of Sac

a.Reference number: S57252; MUID:95397594; PMID:7668046

A.Accession: S57253

A.Molecule type: DNA

A.Residues: 26-859 <COG>

A.Cross-references: EMBL:X83690; NID:g794143; PIDN:CAA58663.1; PID:e224082; PID:g1209471

R.Brusch, C.V.

submitted to the EMBL Data Library, January 1995

C.Accession: T02917; A27863

R.Kunze, R.

submitted to the EMBL Data Library, February 1988

A.Reference number: 214767

A.Accession: T02917

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 1-102 <KUN>

A.Cross-references: UNIPROT:P08771; EMBL:X05424; NID:g22112; PIDN:CAA29006.1; PID:g2211

R.Kunze, R.; Stochaj, U.; Laufs, J.; Starlinger, P.

EMBO J. 6, 1555-1563, 1987

A>Title: Transcription of transposable element Activator (Ac) of Zea mays L.

A.Reference number: A91069

A.Accession: A27863

A.Molecule type: mRNA

A.Residues: 148 <KUZ>

C.Genetics:

A.Mobile element: transposable element Ac

Query Match

Best Local Similarity 90.5%; Score 76; DB 2; Length 102;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12

DB 20 HNNHNNHNNHNN 31

RESULT 3

S71628

Sensory transduction histidine kinase doka - slime mold (Dictyostelium discoideum)

C.Species: Dictyostelium discoideum

C.Date: 27-Nov-1997 #sequence revision 12-Dec-1997 #text change 09-Jul-2004

C.Accession: S71628; S78068

R.Schuster, S.C.; Noegel, A.A.; Oehme, F.; Gerisch, G.; Simon, M.I.

EMBO J. 15, 3880-3889, 1996

A>Title: The hybrid histidine kinase DOKA is part of the osmotic response system of Dict

A.Reference number: S71628; MUID:96324396; PMID:8670893

A.Accession: S71628

A.Molecule type: DNA

A.Residues: 1-1670 <SCH>

A.Cross-references: UNIPROT:Q23901; EMBL:X96869

A.Experimental source: strain AX2; substrain 214

R.Schuster, S.C.; Noegel, A.A.; Oehme, F.; Gerisch, G.; Simon, M.I.

submitted to the EMBL Data Library, March 1996

A.Description: The hybrid histidine kinase DOKA is part of the osmotic response system c

A.Accession: S78068

A.Molecule type: DNA

A.Residues: 1-149, 'E', 151-219, 'TRVLKLIQSTNNWY', 238-1671 <SCW>

A.Cross-references: EMBL:X96869; NID:g1237201; PIDN:CAA65612.1; PID:g1237202

C.Genetics:

A.Function:

A.Description: modulates cell response to changes in osmolarity; involved in spore forma

C.Keywords: phosphoprotein; signal transduction

F;1520-1629/Domain: response regulator homology <RHH2>

F;1568/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match

Best Local Similarity 87.5%; Score 73.5; DB 2; Length 1671;

Matches 12; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 HNNHNN-HNNHNN 12

DB 577 HNNHNNHNNHNN 589

RESULT 4

A46391

cAMP receptor subtype 3 - slime mold (Dictyostelium sp.)

C.Species: Dictyostelium sp.

C/Date: 18-Oct-1993 #sequence_revision 18-Nov-1994 #text_change 29-Oct-1999
C/Accession: A46391
R/Johnson, R.L.; Saxe II, C.L.; Gollop, R.; Kimmel, A.R.; Devreotes, P.N.
Genes Dev. 7, 273-282, 1993
A/Title: Identification and targeted gene disruption of cAR3, a cAMP receptor subtype ex
A/Reference number: A46391; MUID:93170666; PMID:8382181
A/Accession: A46391
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-490 <OH>
A/Cross-references: GB:S55235; NID:g265733; PIDN:AA25437.1; PID:g265734
A/Note: Sequence extracted from NCBI backbone (NCBIN:125367, NCBIPI:125369)
C/Keywords: cAMP binding; G protein-coupled receptor; phosphoprotein; transmembrane prot

Query Match 83.3%; Score 70; DB 2; Length 490;
Best Local Similarity 83.3%; Pred. No. 0.011;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
||||:||||:
DB 325 HNNNNNNNNNN 336

RESULT 5
T49492
hypothetical protein B14D6.440 [imported] - Neurospora crassa
C/Species: Neurospora crassa
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C/Accession: T49492
R/Schulte, U.; Alyn, V.; Hehse, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A/Reference number: Z25022
A/Accession: T49492
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-898 <SCH>
A/Cross-references: EMBL:ALJ56173; GSPDB:GN00116; NCSP:B14D6.440
A/Experimental source: BAC clone B14D6; strain OR74A
C/Genetics:
A/Gene: NCSP:B14D6.440
A/Map position: 6
A/Intons: 818/3

Query Match 79.8%; Score 67; DB 2; Length 898;
Best Local Similarity 83.3%; Pred. No. 0.049;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
||||:||||:
DB 641 NNNNNNNNNNN 652

RESULT 6
AG1396
cation transport protein (efflux) homolog lmo2575 [imported] - Listeria monocytogenes (S
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AG1396
R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Fsihl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltounam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AG1396
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-303 <GLA>
A/Cross-references: UNIPROT:Q9Y480; GB:NC_003210; PIDN:CAD00653.1; PID:g16412063; GSPDB:
A/Experimental source: strain EGD-e
C/Genetics:

A/Gene: lmo2575
C/Superfamily: zinc transporter Znt-2

Query Match 77.4%; Score 65; DB 2; Length 303;
Best Local Similarity 75.0%; Pred. No. 0.03;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
||||:||||:
DB 3 HNNNNNNNNNN 14

RESULT 7
AB1772
cation transport protein (efflux) homolog lin2720 [imported] - Listeria innocua (strain
C/Species: Listeria innocua
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AB1772
R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Fsihl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltounam, A.; M
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AB1772
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-303 <GLA>
A/Cross-references: UNIPROT:Q92788; GB:AL592022; PIDN:CAC97946.1; PID:g16415256; GSPDB:
A/Experimental source: strain C1p11262
C/Genetics:
A/Gene: lin2720
C/Superfamily: zinc transporter Znt-2

Query Match 77.4%; Score 65; DB 2; Length 303;
Best Local Similarity 75.0%; Pred. No. 0.03;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
||||:||||:
DB 3 HNNNNNNNNNN 14

RESULT 8
AB4687
probable homeodomain transcription factor [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: AB4687
R/lin, X.; Kaul, S.; Rounleley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujita, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Taiton, L
eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salberg, S.L.; Frazer, C.M.; Venter,
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: AB4420; MUID:20083467; PMID:10671197
A/Accession: AB4687
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-244 <STO>
A/Cross-references: UNIPROT:Q9SIB4; GB:AE002093; NID:g4580396; PIDN:AAD24374.1; GSPDB:GN
C/Genetics:
A/Gene: At2g28610
A/Map position: 2

Query Match 76.2%; Score 64; DB 2; Length 244;
Best Local Similarity 66.7%; Pred. No. 0.033;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
||||:||||:
DB 115 HNNNNNNNNNN 126

RESULT 9

conserved hypothetical integral membrane protein BB0616 - Lyme disease spirochete
G70176
C/Species: Borrelia burgdorferi (Lyme disease spirochete)
C/Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C/Accession: G70176
R/Frazer, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kervilave, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A/Authors: Smith, H.O.; Venter, J.C.
A/Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A/Reference number: A70100; MUID:98065943; PMID:9403685
A/Accession: G70176
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-451 <KLE>
A/Cross-references: UNIPROT:O51561; GB:AE001163; GB:AE000783; NID:g2688541; PIDN:AC6697
A/Experimental source: strain B31

Query Match

Best Local Similarity 76.2%; Score 64; DB 2; Length 451;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
|:|:|:|:|:|:|:
DB 253 HDHDDHNNHNN 264

RESULT 10

hypothetical protein At2g43970 [imported] - Arabidopsis thaliana
N/Alternative names: hypothetical protein FE13.10
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
R/Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, June 1998
A/Description: Arabidopsis thaliana chromosome II BAC FE13 genomic sequence.
A/Reference number: Z14180
A/Accession: T00677
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-529 <ROU>
A/Cross-references: UNIPROT:O80567; EMBL:AC04005; NID:g3212846; PID:g3212854
A/Experimental source: cultivar Columbia
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
Wu, J.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
Natus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: G84872
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-529 <STO>
A/Cross-references: GB:AE002093; NID:g3212854; PIDN:ACC23405.1; GSPDB:GN00139
A/Genes: FE13.10; At2g43970
A/Map position: 2
A/Intons: 200/3; 228/1; 259/3; 283/3; 310/2; 344/3; 367/3; 421/3

Query Match 76.2%; Score 64; DB 2; Length 529;
Best Local Similarity 75.0%; Pred. No. 0.072;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
|:|:|:|:|:|:|:
DB 458 HNNHNNHNNHNN 469

RESULT 11

hypothetical protein F9D24.30 - Arabidopsis thaliana
T45972
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C/Accession: T45972
R/D'Angelio, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.,
submitted to the Protein Sequence Database, January 2000
A/Reference number: Z23011
A/Accession: T45972
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-329 <DAN>
A/Cross-references: UNIPROT:Q9M2K4; EMBL:AL137081
A/Experimental source: cultivar Columbia; BAC clone F9D24
A/Genetics:
A/Map position: 3
A/Intons: 208/2; 238/3; 281/1
A/Note: F9D24.30

Query Match

Best Local Similarity 73.8%; Score 62; DB 2; Length 329;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
|:|:|:|:|:|:|:
DB 101 HNNHNNHNNHNN 112

RESULT 12

hypothetical protein - fission yeast (Schizosaccharomyces pombe) (fragment)
T43145
C/Species: Schizosaccharomyces pombe
C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 03-Nov-2000
C/Accession: T43145
R/Ishidoka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A/Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A/Reference number: Z17923; MUID:98162722; PMID:9501991
A/Accession: T43145
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-356 <YOS>
A/Cross-references: EMBL:D89239; NID:g1749685; PIDN:BAAL3900.1; PID:g1749686
A/Experimental source: strain PR745
A/Superfamily: zinc transporter Znt-2

Query Match 73.8%; Score 62; DB 2; Length 356;
Best Local Similarity 72.7%; Pred. No. 0.089;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 11
|:|:|:|:|:|:|:
DB 164 HHHHHNNHNDH 174

RESULT 13

hypothetical protein DKFZp564L2123.1 - human (fragment)
T08684
C/Species: Homo sapiens (man)
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C/Accession: T08684
R/Wambutt, R.; Heubner, D.; Mewes, H.W.; Gaassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A/Reference number: Z16469
A/Accession: T08684
A/Molecule type: mRNA
A/Residues: 1-529 <WAM>
A/Cross-references: UNIPROT:Q9Y3Z1; EMBL:AL050294
A/Experimental source: fetal brain; clone DKFZp564L2123
A/Genetics:
A/Note: DKFZp564L2123.1

Query Match 73.8%; Score 62; DB 2; Length 529;
 Best Local Similarity 72.7%; Pred. No. 0.13;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
 |||:|:|:|:|
 DB 124 HSHNHQSHNH 134

RESULT 14

T00119
 Probable transcription factor ken - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster

C/Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004

C/Accession: T00119

R/Lukacsovich, T.; Azeiteiro, Z.; Jun, N.; Awano, W.; Yamamoto, D.

Submitted to the EMBL Data Library, January 1998

A/Description: The Ken and Barbie gene encoding a putative transcription factor with two

A/Reference number: Z14114

A/Accession: T00119

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-601 <LUK>

A/Cross-references: UNIPROT:O7459; EMBL:AB010260; NID:d1228130; PID:d1033648

A/Genes: ken

C/Genetics:

A/Cross-references: FlyBase:FBgn0011236

A/Map position: 2R

C/Keywords: transcription factor; zinc finger

Query Match 73.8%; Score 62; DB 2; Length 601;
 Best Local Similarity 66.7%; Pred. No. 0.15;
 Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
 |||:|:|:|:|
 DB 250 HHHHHHNNNN 261

RESULT 15

T15993
 Hypothetical protein F09E10.9 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000

C/Accession: T15993

R/Geisler, C.; Gattung, S.

Submitted to the EMBL Data Library, November 1995

A/Description: The sequence of C. elegans cosmid F09E10.

A/Reference number: Z18443

A/Accession: T15993

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-126 <GEI>

A/Cross-references: EMBL:U41749; NID:g118144; PID:g118153; PIDN:AAB52490.1; GSPDB:GNOC

A/Experimental source: strain Bristol N2; clone F09E10

C/Genetics:

A/Genes: CESP:F09E10.9

A/Map position: X

A/Intons: 57/3; 96/2

Query Match 72.6%; Score 61; DB 2; Length 126;
 Best Local Similarity 72.7%; Pred. No. 0.042;
 Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
 |||:|:|:|:|
 DB 33 HHHNNHNNH 43

RESULT 16

D64087

Probable peptidylprolyl isomerase (EC 5.2.1.8) Fkbp-type slyD - Haemophilus influenzae

C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004

C/Accession: D64087; T09404

R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kiehlavage,

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman,

; D.M.; Brandon, R.C.; Fink, U.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M.

Science 269, 496-512, 1995

A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A/Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A/Reference number: A64000; NID:95350630; PMID:7542800

A/Accession: D64087

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-190 <TIGR>

A/Cross-references: UNIPROT:P44830; GB:U32753; GB:L42023; NID:g1573701; PIDN:AAC22358.1

C/Genetics:

A/Genes: slyD, H10699

C/Keywords: cis-trans-isomerase

Query Match 72.6%; Score 61; DB 2; Length 190;
 Best Local Similarity 72.7%; Pred. No. 0.064;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
 |||:|:|:|:|
 DB 172 HHHHHHNNNN 182

RESULT 17

A56038
 DNA-binding protein ovo - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster

C/Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004

C/Accession: A56038

R/Garfinkel, M.D.; Wang, J.; Liang, Y.; Mahowald, A.P.

Mol. Cell. Biol. 14, 6809-6818, 1994

A/Title: Multiple products from the shavenbaby-ovo gene region of Drosophila melanogaster

A/Reference number: A56038; MUID:95021209; PMID:7935398

A/Accession: A56038

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-1028 <GAR>

A/Cross-references: UNIPROT:P51521; GB:U11383; NID:g520526; PIDN:AAB60216.1; PID:g52052

C/Genetics:

A/Genes: ovo

A/Cross-references: FlyBase:FBgn0003028

Query Match 72.6%; Score 61; DB 2; Length 1028;
 Best Local Similarity 66.7%; Pred. No. 0.35;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
 |||:|:|:|:|
 DB 144 HHHHHHNNNN 155

RESULT 18

S16356
 ovo protein - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster

C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C/Accession: S16356

R/Mevel-Ninio, M.; Terracol, R.; Kafatos, F.C.

EMBO J. 10, 2259-2266, 1991

A/Title: The ovo gene of Drosophila encodes a zinc finger protein required for female ge

A/Reference number: S16356; MUID:91293102; PMID:1712294

A/Accession: S16356

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1213 <MEV>

A/Cross-references: UNIPROT:Q8T8L9; EMBL:X59772

C/Genetics:

A/Genes: FlyBase:ovo

A/Cross-references: FlyBase:FBgn0003028

A:Introns: 931/3; 1152/3

Query Match 72.6%; Score 61; DB 2; Length 1213;
Best Local Similarity 66.7%; Pred. No. 0.42;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:|:|
DB 507 HNNHHHHQHNNN 518

RESULT 19

B97242

hypothetical protein CAC2781 [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004

C:Accession: B97242

R:Noiling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld

A:Reference number: A96900; PMID:21359325; PMID:21359325

A:Accession: B97242

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-180 <KUR>

A:Cross-references: UNIPROT:Q97FP9; GB:AE001437; PDB:AAK80725.1; PDB:G15025820; GSPDB:G

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetic:

A:Gene: CAC2781

Query Match

Best Local Similarity 70.2%; Score 59; DB 2; Length 180;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
|:|:|:|:|:|:|
DB 31 HNNHNNHNNH 41

RESULT 20

G97048

probable membrane protein [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004

C:Accession: G97048

R:Noiling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld

A:Reference number: A96900; PMID:21359325; PMID:21359325

A:Accession: G97048

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-200 <KUR>

A:Cross-references: UNIPROT:Q97JH7; GB:AE001437; PDB:AAK79178.1; PDB:G15024129; GSPDB:G

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetic:

A:Gene: CAC1206

Query Match

Best Local Similarity 70.2%; Score 59; DB 2; Length 200;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
|:|:|:|:|:|:|
DB 117 HNNHNNHNNH 127

RESULT 21

B64216

hypothetical protein MG149 - Mycoplasma genitalium
C:Species: Mycoplasma genitalium

C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

C:Accession: B64216

R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.; Fink, M.; Fullman, J.; Nguyen, D.; Uterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.

A:Title: The minimal gene complement of Mycoplasma genitalium.

A:Reference number: A64200; PMID:96026346; PMID:7569993

A:Accession: B64216

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-281 <TIGR>

A:Cross-references: UNIPROT:P47395; GB:U39694; GB:U43967; NID:G1045822; PDB:G1045832; T

A:Experimental source: strain G-37

C:Genetic:

A:Gene code: SGC3

C:Superfamily: Mycoplasma pneumoniae probable lipoprotein VxpSP7_orf320

Query Match 70.2%; Score 59; DB 2; Length 281;
Best Local Similarity 72.7%; Pred. No. 0.18;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
|:|:|:|:|:|:|
DB 146 HNNHNNHNNH 156

RESULT 22

A46390

CAMP receptor subtype 2, CAR2 - slime mold (Dictyostelium discoideum)

C:Species: Dictyostelium discoideum

C:Date: 18-Oct-1993 #sequence_revision 18-Nov-1994 #text_change 29-Oct-1999

C:Accession: A46390

R:Saxe III, C.L.; Ginsburg, G.T.; Louis, J.M.; Johnson, R.; Devreotes, P.N.; Kimmel, A.R.

A:Title: CAR2, a prestalk CAMP receptor required for normal tip formation and late devel.

A:Reference number: A46390; PMID:9310665; PMID:8436297

A:Accession: A46390

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-375 <SNX>

A:Cross-references: GB:555234; NID:9265731; PDB:AAZ5436.1; PDB:9265732

A:Note: Sequence inconsistent with the nucleotide translation

A:KeyWords: CAMP binding; G protein-coupled receptor

Query Match 70.2%; Score 59; DB 2; Length 375;
Best Local Similarity 58.3%; Pred. No. 0.24;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:|:|
DB 346 HNNHNNHNNH 357

RESULT 23

C84905

probable extensin [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C:Accession: C84905

R:Jin, X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Benito, M.T.; Town, C.D.; Fujii, C.Y.;

enes, D.; Niernan, M.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; PMID:20083487; PMID:1061197

A:Accession: C84905

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-394 <STO>

A:Cross-references: UNIPROT:Q92NJJ; GB:AE002093; NID:93831447; PDB:AAK69330.1; GSPDB:GN

C:Genetic:

A:Gene: At2g46630
A:Map position: 2

Query Match
Best Local Similarity 70.2%; Score 59; DB 2; Length 394;
Matches 8; Conservativity 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
|:|:|:|:|:|:
DB 217 HNNHNNHNNHNN 228

RESULT 24

151436
HNF-3beta - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C/Accession: 151436; D56556
R/Ruitz I Altaba, A.; Prezioso, V.R.; Darnell, J.E.; Jessell, T.M.
Mech. Dev. 44, 91-108, 1993
A>Title: Sequential expression of HNF-3 beta and HNF-3 alpha by embryonic organizing center
A/Reference number: 151436; PMID:94206853; PMID:8155584
A/Accession: 151436
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-434 <RUI>
A/Cross-references: UNIPROT:Q91765; GB:U25637; NID:9409773; PIDN:AAA20679.1; PID:9409774
R/Knochel, S.; Lef, J.; Clement, J.; Klocke, B.; Hille, S.; Koster, M.; Knochel, W.
Mech. Dev. 38, 157-165, 1992
A>Title: Activin A induced expression of a fork head related gene in posterior chordamesoderm
A/Reference number: A56556; PMID:93041288; PMID:1358174
A/Accession: D56556
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 140-250 <KNO>
A/Experimental source: gastrula
A/Note: sequence extracted from NCBI backbone (NCBIP:118180)
C/Superfamily: transcription factor HNF-3; fork head DNA-binding domain homology
C/Keywords: DNA binding; transcription factor
F:149-240/Domain: fork head DNA-binding domain homology <FHD>

Query Match
Best Local Similarity 70.2%; Score 59; DB 2; Length 434;
Matches 7; Conservativity 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
|:|:|:|:|:|:
DB 363 HNNHNNHNNHNN 374

RESULT 25

F96836
Hypothetical protein T21F11.27 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: F96836
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chen, C.R.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ameen, N.R.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maitl, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; PMID:21016719; PMID:11130712
A/Accession: F96836
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-444 <STO>
A/Cross-references: UNIPROT:Q9W8L6; GB:AE005173; NID:96730739; PIDN:AAF27129.1; GSPDB:GN
C/Genetics:

A:Gene: T21F11.27
A:Map position: 1

Query Match
Best Local Similarity 70.2%; Score 59; DB 2; Length 444;
Matches 7; Conservativity 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
|:|:|:|:|:|:
DB 312 HNNHNNHNNHNN 323

RESULT 26

T46024
Hypothetical protein T10K17.230 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C/Accession: T46024
R/Benes, V.; Wurmach, E.; Drzonek, H.; Ansgorge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.J.
submitted to the Protein Sequence Database, January 2000
A/Reference number: Z23019
A/Accession: T46024
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-580 <BEN>
A/Cross-references: UNIPROT:Q9M2P6; EMBL:AL132977
A/Experimental source: cultivar Columbia; BAC clone T10K17
C/Genetics:
A/Map position: 3
A/Intons: 81/2; 224/2; 253/3; 280/2; 379/2; 397/3; 456/1
A/Note: T10K17.230

Query Match
Best Local Similarity 70.2%; Score 59; DB 2; Length 580;
Matches 7; Conservativity 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
|:|:|:|:|:|:
DB 529 HNNHNNHNNHNN 540

RESULT 27

T13389
Hypothetical protein 115C2.10 - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C/Accession: T13389
R/Salles, C.; Valenti, P.; Darlamitsou, A.; Henderson, N.; Campbell, L.; Glover, D.
submitted to the EMBL Data Library, May 1999
A/Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A/Reference number: Z17665
A/Accession: T13389
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1291 <CAT>
A/Cross-references: UNIPROT:O77261; EMBL:AL031581; NID:e1320978; PID:e1426292; PIDN:CAM.
C/Genetics:
A/Cross-references: FlyBase:FBgn0020381
A/Map position: X
A/Intons: 238/3; 1225/1
A/Note: EG:115C2.10

Query Match
Best Local Similarity 70.2%; Score 59; DB 2; Length 1291;
Matches 7; Conservativity 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
|:|:|:|:|:|:
DB 544 HNNHNNHNNHNN 555

RESULT 28

S61046

A/Cross-references: UNIPROT:Q9T0D9; EMBL:AL049500
A/Experimental source: cultivar Columbia; BAC clone T5C23
C/Genetics:
A/Map position: 4
A/Intons: 143/3; 173/2; 217/1; 237/2; 273/3; 318/3; 352/2; 401/3; 412/3; 444/1; 473/3;
A/Note: T5C23.150

Query Match 67.9%; Score 57; DB 2; Length 658;
Best Local Similarity 58.3%; Pred. No. 0.77;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNNNNNNNH 12
|:|:|:|:|:
DB 571 HHHHKKRHHN 582

RESULT 33
S4615
hypothetical protein YBR238c - yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein YBR1608
C/Species: Saccharomyces cerevisiae
C/Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C/Accession: S4615
R/Aljnovic, G.; Pohl, F.M.; Pohl, T.M.
submitted to the Protein Sequence Database, August 1994
A/Reference number: S45906
A/Accession: S4615
A/Molecule type: DNA
A/Residues: 1731 <ALJ>
A/Cross-references: UNIPROT:P38330; EMBL:Z36107; NID:G536642; PIDN:CAA85201.1; PID:G5366
A/Experimental source: strain S288C
C/Genetics:
A/Gene: MIPS:YBR238c
A/Cross-references: SGD:S0000442
A/Map position: 2R
C/Superfamily: hypothetical protein YBR238c

Query Match 67.9%; Score 57; DB 2; Length 721;
Best Local Similarity 66.7%; Pred. No. 0.85;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNNNNNNNH 12
|:|:|:|:|:
DB 100 HNNNNNNNNH 111

RESULT 34
H86642
hypothetical protein yhec [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C/Species: Lactococcus lactis subsp. lactis
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C/Accession: H86642
R/Bolotin, A.; Winkler, P.; Mauger, S.; Jallou, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A/Reference number: A86625; MUID:21235186; PMID:11337471
A/Accession: H86642
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-107 <STO>
A/Cross-references: UNIPROT:Q9CJ62; GB:AE005176; PID:G12722993; PIDN:AAK04242.1; GSPDB:C
A/Experimental source: strain IL1403
C/Genetics:
A/Gene: yhec

Query Match 66.7%; Score 56; DB 2; Length 107;
Best Local Similarity 63.6%; Pred. No. 0.17;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNNNNNNNH 11
|:|:|:|:|:
DB 3 HTDHDHHDH 13

RESULT 35
S49807
hypothetical protein YML053c - yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein YN958.09c
C/Species: Saccharomyces cerevisiae
C/Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C/Accession: S49807
R/Devlin, K.; Churcher, C.
submitted to the EMBL Data Library, November 1994
A/Reference number: S49800
A/Accession: S49807
A/Molecule type: DNA
A/Residues: 1-212 <DEV>
A/Cross-references: UNIPROT:Q04978; EMBL:Z46729; NID:G577134; PIDN:CAA86722.1; PID:G577
C/Genetics:
A/Gene: MIPS:YML053c
A/Cross-references: SGD:S0004517
A/Map position: 13L

Query Match 66.7%; Score 56; DB 2; Length 212;
Best Local Similarity 66.7%; Pred. No. 0.33;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNH 12
|:|:|:|:|:
DB 90 NNNNNNNNNH 101

RESULT 36
T19585
hypothetical protein C30H6.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T19585
R/Mortimore, B.
submitted to the EMBL Data Library, October 1996
A/Reference number: Z19148
A/Accession: T19585
A/Status: preliminary; translated from GB/EMBL/DDBT
A/Molecule type: DNA
A/Residues: 1-360 <WIL>
A/Cross-references: UNIPROT:Q9XVR4; EMBL:Z81044; PIDN:CA802806.1; GSPDB:GN00022; CESP:C
A/Experimental source: clone C30H6
C/Genetics:
A/Gene: CESP:C30H6.2
A/Map position: 4
A/Intons: 43/2; 67/3; 111/3; 149/3; 231/3
C/Superfamily: Caenorhabditis elegans hypothetical protein C30H6.2

Query Match 66.7%; Score 56; DB 2; Length 360;
Best Local Similarity 63.6%; Pred. No. 0.56;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNNNNNNNH 11
|:|:|:|:|:
DB 72 HDGHNHGHSH 82

RESULT 37
A84696
probable zinc transporter [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: A84696
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
ues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: A84696
A/Status: preliminary

A:Molecule type: DNA
A:Residues: 1-385 <STO>
A:Cross-references: UNIPROT:Q9ZM23; GB:AE002093; NID:g3980394; PIDN:AA095197.1; GSPDB:GN
A:Gene: At2g29410
A:Map position: 2

Query Match 66.7%; Score 56; DB 2; Length 385;
Best Local Similarity 63.6%; Pred. No. 0.6;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
|:|:|:|:|:|
Db 202 HHHHHHHKHQ 212

RESULT 38
S19365

hypothetical protein YCL037c - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C/Accession: S19365
R:Delgado, M.; Betebedan, M.; Navas, L.
Submitted to the Protein Sequence Database, March 1992
A:Reference number: S19365
A:Accession: S19365
A:Molecule type: DNA
A:Residues: 1-466
A:Cross-references: UNIPROT:P25567; EMBL:X59720; NID:g1907116; PIDN:CAA42379.1; PID:g5332
A:Gene: SGD:SRO9
A:Cross-references: SGD:S0000542; MIPS:YCL037c
A:Map position: 3L

Query Match 66.7%; Score 56; DB 2; Length 466;
Best Local Similarity 58.3%; Pred. No. 0.73;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:|
Db 206 HHRNHHSHHNN 217

RESULT 39
T18437

hypothetical protein C0405c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T18437
R:Lawson, D.; Bowman, S.; Barrell, B.
Submitted to the EMBL Data Library, August 1997
A:Reference number: Z18935
A:Accession: T18437
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-608 <LAW>
A:Cross-references: UNIPROT:O77331; EMBL:Z98547; NID:e1325376; PID:e1325391; PIDN:CAB111
A:Gene: C0405c
A:Map position: 3

Query Match 66.7%; Score 56; DB 2; Length 608;
Best Local Similarity 66.7%; Pred. No. 0.96;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:|
Db 413 HNNNNNNNNNN 424

RESULT 40
C88402
protein H05C05.1 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: C88402
R:anonymus, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998

A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; PMID:99069613; PMID:9851916
A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A:Accession: C88402
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-828 <STO>
A:Cross-references: UNIPROT:Q9TXU2; GB:chr_III; PIDN:AA068790.1; PID:g3790735; GSPDB:GN
A:Gene: H05C05.1
A:Map position: 3

Query Match 66.7%; Score 56; DB 2; Length 828;
Best Local Similarity 58.3%; Pred. No. 1.3;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:|
Db 568 HHRSHHHHNNH 579

RESULT 41
T18429

hypothetical protein C0345w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T18429
R:Lawson, D.; Bowman, S.; Barrell, B.
Submitted to the EMBL Data Library, August 1997
A:Reference number: Z18935
A:Accession: T18429
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1711 <LAW>
A:Cross-references: UNIPROT:O77332; EMBL:Z98547; NID:e1325376; PID:e1325381; PIDN:CAB111
A:Gene: C0345w
A:Note: PFC0345w

Query Match 66.7%; Score 56; DB 2; Length 1711;
Best Local Similarity 66.7%; Pred. No. 2.7;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:|
Db 1420 HNNNNNNNNNN 1431

RESULT 42
A47327

selenoprotein P precursor [validated] - human
C:Species: Homo sapiens (man)
C>Date: 21-Sep-1993 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C/Accession: A47327; S42752
R:Hill, K.E.; Llovd, R.S.; Burk, R.F.
Proc. Natl. Acad. Sci. U.S.A. 90, 537-541, 1993
A>Title: Conserved nucleotide sequences in the open reading frame and 3' untranslated re
A:Reference number: A47327; PMID:93133823; PMID:8421687
A:Accession: A47327
A:Molecule type: mRNA
A:Residues: 1-381 <HIL>
A:Cross-references: UNIPROT:P49908; GB:Z11793; NID:g36425; PIDN:CAA77836.1; PID:g2654365
A:Experimental source: heart and liver
A:Note: in Genbank entry HSELEPM, release 117.0, PIDN:CAA77836.1, the selenocysteine UGA
Biochim. Biophys. Acta 1204, 243-249, 1994
A>Title: Purification of selenoprotein P from human plasma.
A:Reference number: S42752; PMID:94191007; PMID:8142465

A:Accession: S42752
A:Molecule type: Protein
A:Residues: 20-27, 'X', 29-33 <AKE>
A>Note: mature forms of 55k and 61k were detected in plasma; the protein was shown to co
C:Genetics:
A:Gene: GDB:SEPP1; SLNP
A:Cross-references: GDB:138278; OMIM:601484
A:Map position: 5q31-5q31
C:Function:
A:Description: may act as a free-radical scavenger
C:Superfamily: selenoprotein P
C:Keywords: extracellular protein; glycoprotein; heparin binding; liver; plasma; seleno
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-381/Product: selenoprotein P #status experimental <MAT>
F:46,83,119,128,338/Binding site: carboxylate (Asn) (covalent) #status predicted
F:59,300,318,330,345,352,367,369,376,378/Modified site: selenocysteine #status predicted

Query Match 65.5%; Score 55; DB 1; Length 381;
Best Local Similarity 63.6%; Pred. No. 0.81;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 HNNHHNNHHN 11
DB 207 HHHHHHHHHH 217

RESULT 43
A96625
hypothetical protein T2K10.11 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A96625
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.; Li, J.H.; Li, Y.; Lim, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani,
Rizzo, M.; Rooney, T.J.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; PMID:21016719; PMID:11130712
A:Accession: A96625
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1386 <STO>
A:Cross-references: UNIPROT:Q9ZU17; GB:AE005173; NID:94249385; PIDN:AD14482.1; GSPDB:GN
A:Genetics:
A:Gene: T2K10.11
A:Map position: 1

Query Match 65.5%; Score 55; DB 2; Length 386;
Best Local Similarity 66.7%; Pred. No. 0.82;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHHNNHHN 12
DB 346 HNNHHNNHHN 357

RESULT 44
S61568
probable membrane protein YDR205w - yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein YD8142A.02
C:Species: Saccharomyces cerevisiae
C>Date: 09-Mar-1996 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004
C:Accession: S61568
R:Murphy, L.; Harris, D.
submitted to the EMBL Data Library, December 1995
A:Reference number: S61117
A:Accession: S61568
A:Molecule type: DNA

A:Residues: 1-740 <MUR>
A:Cross-references: UNIPROT:Q03455; EMBL:Z68194; NID:91204148; PID:e213788; PID:g112233
A:Experimental source: strain AB972
C:Genetics:
A:Gene: SGD:MSC2; MIPS:YDR205w
A:Cross-references: SGD:S0002613
A:Map position: 4R
C:Keywords: transmembrane protein
F:27-38/Domain: transmembrane #status predicted <TM1>
F:107-123/Domain: transmembrane #status predicted <TM2>
F:155-171/Domain: transmembrane #status predicted <TM3>
F:156-112/Domain: transmembrane #status predicted <TM4>
F:219-255/Domain: transmembrane #status predicted <TM5>
F:264-280/Domain: transmembrane #status predicted <TM6>
F:318-334/Domain: transmembrane #status predicted <TM7>
F:346-362/Domain: transmembrane #status predicted <TM8>
F:404-420/Domain: transmembrane #status predicted <TM9>
F:437-453/Domain: transmembrane #status predicted <TM10>
F:477-493/Domain: transmembrane #status predicted <TM11>
F:508-524/Domain: transmembrane #status predicted <TM12>
F:544-560/Domain: transmembrane #status predicted <TM13>
F:576-592/Domain: transmembrane #status predicted <TM14>

Query Match 65.5%; Score 55; DB 2; Length 740;
Best Local Similarity 63.6%; Pred. No. 1.6;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHHNNHHN 11
DB 369 HSHHHHHHHH 379

RESULT 45
T18469
hypothetical protein C0455w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18469
R:Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z18937
A:Accession: T18469
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-786 <LAW>
A:Cross-references: UNIPROT:O77357; EMBL:AL008970; NID:61407852; PID:e1332563; PIDN:CAA
A:Genetics:
A:Map position: 3
A>Note: C0455w

Query Match 65.5%; Score 55; DB 2; Length 786;
Best Local Similarity 66.7%; Pred. No. 1.7;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHHNNHHN 12
DB 216 HAHNNHHNNN 227

RESULT 46
T51065
hypothetical protein B12F1.170 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 08-Sep-2000
C:Accession: T51065
R/Schulte, U.; Algn, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T51065
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-172 <SCH>
A:Cross-references: EMBL:AL390091; GSPDB:GN00116; NCSP:B12F1.170

A:Experimental source: BAC clone B12F1, strain OR74A
C:Genetics:
A:Gene: NCSP:B12F1.170
A:Map position: 6
A:Introns: 51/3: 125/2
C:Superfamily: Neurospora crassa hypothetical protein B12F1.170

Query Match 64.3%; Score 54; DB 2; Length 172;
Best Local Similarity 58.3%; Pred. No. 0.49;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 HNNHNNHNNH 12
DB 87 HHHHHNNHHOH 98

RESULT 47
B49122
Brn-3b protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 25-Dec-1994 #sequence_revision 01-Sep-1995 #text_change 16-Aug-2004
C:Accession: S41512
R:Title: T.; McLean-Hunter, S.; Zoernig, M.; Moercoy, T.
Nucleic Acids Res. 21, 5921-5929, 1993
A:Title: Mouse Brn-3 family of POU transcription factors: a new aminoterminal domain is
A:Reference number: S41511; MUID:94119691; PMID:8290353
A:Accession: S41512
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-236 <THE>
C:Superfamily: homeobox homology; POU domain homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:180-236/Domain: homeobox homology <HOX>

Query Match 64.3%; Score 54; DB 2; Length 236;
Best Local Similarity 50.0%; Pred. No. 0.68;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
QY 1 HNNHNNHNNH 12
DB 78 HHHHHNNHHHH 89

RESULT 48
A56837
homeotic protein MOX2 - human
C:Species: Homo sapiens (man)
C>Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 09-Jul-2004
C:Accession: A56837; S52168
R:Grigoriou, M.; Kastirnak, M.C.; Modi, W.S.; Theodorakis, K.; Mankoo, B.; Pachnis, V.;
Genomics 26, 550-555, 1995
A:Title: Isolation of the human MOX2 homeobox gene and localization to chromosome 7p22.1
A:Reference number: A56837; MUID:95311791; PMID:7607679
A:Accession: A56837
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-303 <GRI>
C:Cross-references: UNIPROT:P50222, EMBL:X82629; NID:G732790; PIDN:CAAS7949.1; PID:G7327
R:Grigoriou, M.; Kastirnak, M.C.; Modi, W.; Theodorakis, K.; Mankoo, B.; Pachnis, V.;
submitted to the EMBL Data Library, November 1994
A:Description: Isolation of the human MOX2 homeobox gene and localization to chromosome
A:Reference number: S52168
A:Accession: S52168
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-7, 'R', '9-110, 'R', '112-122, 'Q', '124-157, 'V', '159-303 <GR2>
C:Cross-references: EMBL:X82629
C:Genetics:
A:Gene: GDB:MOX2
A:Cross-references: GDB:120703; OMIM:155970
A:Map position: 3q13-q13
C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:187-243/Domain: homeobox homology <HOX>

Query Match 64.3%; Score 54; DB 2; Length 303;
Best Local Similarity 50.0%; Pred. No. 0.87;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
QY 1 HNNHNNHNNH 12
DB 68 HHHHHNNHHHH 79

RESULT 49
B49122
homeobox protein Mox-2 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 19-Dec-1993 #sequence_revision 17-Nov-1995 #text_change 16-Aug-2004
C:Accession: B49122; S41779; S29902
R:Candia, A.F.; Hu, J.; Crosby, J.; Lally, P.A.; Noden, D.; Nadeau, J.H.; Wright, C.V.
Development 116, 1123-1136, 1992
A:Title: Mox-1 and Mox-2 define a novel homeobox gene subfamily and are differentially
A:Reference number: A49122; MUID:93201999; PMID:11363541
A:Accession: B49122
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-303 <CAN1>
C:Cross-references: UNIPROT:P32443; EMBL:Z16406; NID:G57949; PIDN:CAA78899.1; PID:G579
A:Experimental source: C57BL
A>Note: the complete translation is not shown
R:Candia, A.F.; Kovalik, J.P.; Wright, C.V.E.
Nucleic Acids Res. 21, 4982, 1993
A:Title: Amino acid sequence of Mox-2 and comparison to its Xenopus and rat homologs.
A:Reference number: S41669; MUID:94232829; PMID:7909944
A:Accession: S41779
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-303 <CAN2>
C:Cross-references: EMBL:Z16406; NID:G57949; PIDN:CAA78899.1; PID:G57950
R:Candia, A.;
submitted to the EMBL Data Library, October 1992
A:Reference number: S29902
A:Accession: S29902
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 186-246 <CAN>
C:Cross-references: EMBL:Z16406
C:Superfamily: homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:187-243/Domain: homeobox homology <HOX>

Query Match 64.3%; Score 54; DB 2; Length 303;
Best Local Similarity 50.0%; Pred. No. 0.87;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
QY 1 HNNHNNHNNH 12
DB 68 HHHHHNNHHHH 79

RESULT 50
A48130
growth arrest-specific homeobox protein Cvx - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A48130; S31976
R:Gorfski, D.H.; LePage, D.F.; Patel, C.V.; Copeland, N.G.; Jenkins, N.A.; Walsh, K.
Mol. Cell. Biol. 13, 3722-3733, 1993
A:Title: Molecular cloning of a diverged homeobox gene that is rapidly down-regulated du
A:Reference number: A48130; MUID:93268321; PMID:8098844
A:Accession: A48130
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-303 <GOR>

A:Cross-references: UNIPROT:P39020; GB:E17223; NID:g57951; PID:CAA78931.1; PID:g57952
A:Experimental source: aorta
A:Note: sequence extracted from NCBI backbone (NCBIN:132842, NCBIP:132843)
R:Coraki, D.H.; LePage, D.F.; Patel, C.V.; Copeland, N.G.; Jenkins, N.A.; Walsh, K.
submitted to the EMBL Data Library, October 1992
A:Description: Cvx: A diverged homeodomain gene which is rapidly down-regulated following
F:187-243/Domain: homeobox homology <Hox>
A:Reference number: S31976
A:Accession: S31976
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-182, 'S', 184-303 <G02>
A:Cross-references: EMBL:Z17223
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:187-243/Domain: homeobox homology <Hox>
Query Match 64.3%; Score 54; DB 2; Length 303;
Best Local Similarity 50.0%; Pred. No. 0.87;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
OY 1 HNNHNNHNNHNN 12
Db 68 HHHHHHHHHHHH 79
RESULT 51
T44684
hypothetical protein Cbl [imported] - Bacillus megaterium
C:Species: Bacillus megaterium
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #ext_change 12-Jul-2004
C:Accession: T44684
R:Rauch, E.; Lancis, A.; Warren, M.J.; Rambach, A.; Thernas, C.
Biochem. J. 335, 159-166, 1998
A:Title: Cobalamin (vitamin B12) biosynthesis: identification and characterization of a
A:Reference number: Z22829; MUID:98416126; PMID:9742225
A:Accession: T44684
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-306 <RAU>
A:Cross-references: UNIPROT:O87690; EMBL:AJ000758; NID:g3724036; PID:CAM04308.1; PID:g3
A:Experimental source: strain DSM 509
C:Genetics:
A:Note: cbix
C:Superfamily: cobaltochelatase/ferrochelatase Cbix/StirB
Query Match 64.3%; Score 54; DB 2; Length 306;
Best Local Similarity 50.0%; Pred. No. 0.88;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
OY 1 HNNHNNHNNHNN 12
Db 274 HHHHHHHHHHHH 285
RESULT 52
E81923
probable periplasmic binding protein NMA0789 [imported] - Neisseria meningitidis (strain
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #ext_change 09-Jul-2004
C:Accession: E81923
R:Parhill, J.; Achman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: E81923
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-308 <PAR>
A:Cross-references: UNIPROT:Q9JVL4; GB:AL162754; GB:AL157959; NID:g3739424; PID:CAM8407
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA0788; NMA0789

C:Superfamily: adhesion B
Query Match 64.3%; Score 54; DB 2; Length 308;
Best Local Similarity 50.0%; Pred. No. 0.89;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
OY 1 HNNHNNHNNHNN 12
Db 115 HHHHHHHHHHHH 126
RESULT 53
D83483
probable metal transporter PA1297 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #ext_change 09-Jul-2004
C:Accession: D83483
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path.
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83483
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-326 <STO>
A:Cross-references: UNIPROT:Q91447; GB:AE004559; GB:AE004091; NID:g9947228; PID:AMG046
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1297
C:Superfamily: zinc transporter Znt-2
Query Match 64.3%; Score 54; DB 2; Length 326;
Best Local Similarity 50.0%; Pred. No. 0.94;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
OY 1 HNNHNNHNNHNN 12
Db 160 HHHHHHHHHHHH 171
RESULT 54
E95858
conserved hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) magap1
C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #ext_change 09-Jul-2004
C:Accession: E95858
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernal
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95858
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <KUR>
A:Cross-references: UNIPROT:Q92X29; GB:AL591985; PID:CA64853.1; PID:g15140005; GSPDB:
A:Experimental source: strain 1021, megaplasmid pSymB
R:Gallbert, F.; Finan, T.M.; Long, S.R.; Punler, A.; Abola, P.; Ampe, F.; Barloy-Hudler,
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Käss, E.; Komp, C.; Lelaure,
beault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Smb20133
A:genome: plasmid
C:Superfamily: conserved hypothetical protein yhcI
Query Match 64.3%; Score 54; DB 2; Length 349;

Best Local Similarity 50.0%; Pred. No. 1;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:
Db 230 HHHHHHHHHHH 241

RESULT 55
H86187
hypothetical protein [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: H86187
R/Title: H86187
A/Authors: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: H86187
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-391 <STO>
A/Cross-references: UNIPROT:Q23038; GB:AE005172; NID:g288565; PIDN:AA671446.1; GSPDB:GN
C/Genetics:
A/Map position: 1
C/Superfamily: pectinesterase pemb

Query Match
Best Local Similarity 50.0%; Pred. No. 1.1;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:
Db 47 HHHHHHHHHHH 58

RESULT 56
G84783
probable pectinesterase [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: G84783
R/Title: H84783
A/Authors: Rounsaey, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; euse, D.; Niemman, W.C.; White, O.; Eiken, J.A.; Salberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: G84783
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-407 <STO>
A/Cross-references: UNIPROT:Q92Q43; GB:AE002093; NID:g4415916; PIDN:AA20147.1; GSPDB:GN
C/Genetics:
A/Map position: 2

Query Match
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:
Db 52 HHHHHHHHHHH 63

RESULT 57
I38502

gene Brn-3b protein - human
C/Species: Homo sapiens (man)
C/Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 16-Aug-2004
C/Accession: I38502; S34121; I59424
R/Title: H38502; S34121; I59424
Neuron 11, 689-701, 1993
A/Title: Brn-3b: a POU domain gene expressed in a subset of retinal ganglion cells.
A/Reference number: I38502; MUID:94000832; PMID:7691107
A/Accession: I38502
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-410 <RES>
A/Cross-references: UNIPROT:Q12837; EMBL:U06233; NID:g458390; PIDN:AA16509.1; PID:g4583
R/Title: The human Brn-3b POU domain transcription factor shows only limited homology to the Br
Nucleic Acids Res. 21, 2946, 1993
A/Title: The human Brn-3b POU domain transcription factor shows only limited homology to the Br
A/Accession: S34121; MUID:93324388; PMID:8332509
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 146-165; S'167-182, 184-410 <RIN>
A/Cross-references: EMBL:X71488; NID:9312468; PIDN:CAA50589.1; PID:g312469
R/Title: Acad. Sci. U.S.A. 93, 10260-10264, 1994
Proc. Natl. Acad. Sci. U.S.A. 93, 10260-10264, 1994
A/Title: Differential expression of four members of the POU family of proteins in activa
A/Reference number: I59424
A/Accession: I59424
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 275-392 <RE2>
A/Cross-references: GB:I20434; NID:g435995; PIDN:AAA6393.1; PID:g508494
A/Genes: Brn-3b
C/Superfamily: homeobox homology; POU domain homology
C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
F/347-403/Domain: homeobox homology <HOK>

Query Match
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:
Db 172 HHHHHHHHHHH 183

RESULT 58
I58156
Brn-3.2 - mouse
C/Species: Mus sp. (mouse)
C/Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 16-Aug-2004
C/Accession: I58156
R/Title: H58156
Neuron 12, 205-218, 1994
A/Title: Brn-3.2: a Brn-3-related transcription factor with distinctive central nervous
A/Reference number: I58156; MUID:94121912; PMID:7904822
A/Accession: I58156
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-411 <RES>
A/Cross-references: GB:S68377; NID:g545068; PIDN:AA60672.1; PID:g545069
C/Genetics:
A/Genes: Brn-3.2
C/Superfamily: homeobox homology; POU domain homology
C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
F/348-404/Domain: homeobox homology <HOK>

Query Match
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
QY 1 HNNHNNHNNHNN 12
|:|:|:|:|:|:
Db 173 HHHHHHHHHHHH 184

RESULT 59

A48273

delta/Yyl/NF-E1/UCRPB transcription factor - mouse

C/Species: Mus musculus (house mouse)

C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004

C/Accession: A48273; A42055; A56418

R/Safirany, G.; Perry, R.P.

Proc. Natl. Acad. Sci. U.S.A. 90, 5559-5563, 1993

A/Title: Characterization of the mouse gene that encodes the delta/Yyl/NF-E1/UCRPB trans

A/Reference number: A48273; MUID:93296177; PMID:8516301

A/Accession: A48273

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-414 <RES>

A/Cross-references: UNIPROT:Q00899; GB:U13968; NID:g293847; PIDN:AAA40477.1; PID:g293845

R/Plangan, J.R.; Becker, K.G.; Emiet, D.L.; Gleason, S.L.; Driggers, P.H.; Levi, B.Z.;

Mol. Cell. Biol. 12, 38-44, 1992

A/Title: Cloning of a negative transcription factor that binds to the upstream conserved

A/Reference number: A42055; MUID:92107191; PMID:1309593

A/Accession: A42055

A/Molecule type: mRNA

A/Residues: 1-414 <FLA>

A/Cross-references: GB:W73963; NID:g202270; PIDN:AAA40522.1; PID:g202271

A/Note: sequence extracted from NCBI backbone (NCBIN:74641, NCBI:P:74642)

R/Harlharan, N.; Kelley, D.E.; Perry, R.P.

Proc. Natl. Acad. Sci. U.S.A. 88, 9799-9803, 1991

A/Title: delta, a transcription factor that binds to downstream elements in several poly

A/Reference number: A56418; MUID:92052178; PMID:1946404

A/Accession: A56418

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-218 'S', 220-374 'G', 376-414 <HAR>

A/Cross-references: GB:M74590; NID:g192940; PIDN:AAA37521.1; PID:g192941

C/Genetics: 227/1; 281/2; 301/3; 354/3

A/Introns: 2

C/Keywords: transcription factor; zinc finger

Query Match Best Local Similarity 64.3%; Score 54; DB 2; Length 414;

Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
|:|:|:|:|:|:
Db 71 HHHHHHHHHHHH 82

RESULT 60

T39712

hypothetical protein SPBC17D11.01 - fission yeast (Schizosaccharomyces pombe)

N/Alternate names: hypothetical protein SPBC20F10.11

C/Species: Schizosaccharomyces pombe

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-Jan-2000

C/Accession: T39712; T39882

R/Beck, A.; Reinhardt, R.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, September 1998

A/Reference number: Z21872

A/Accession: T39712

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-420 <BEC>

A/Cross-references: EMBL:AL031739; PIDN:CAA21072.1; GSPDB:GN00067; SPDB:SPBC17D11.01

R/Lyne, M.; McDougall, R.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.

submitted to the EMBL Data Library, February 1998

A/Reference number: Z21888

A/Accession: T39882

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-281 <LYN>
A/Cross-references: EMBL:AL021747; PIDN:CAA16851.1; GSPDB:GN00067
C/Genetics: 2
A/Map position: 2
A/Introns: 10/1; 28/2
C/Superfamily: hypothetical protein SPBC17D11.01

Query Match Best Local Similarity 64.3%; Score 54; DB 2; Length 420;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
|:|:|:|:|:|:
Db 299 HHHHHHHHHHHH 310

RESULT 61

T42516

hypothetical protein - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe

C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000

C/Accession: T42516

R/Yoshioke, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.

DNA Res. 4, 363-369, 1997

A/Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.

A/Reference number: Z17323; MUID:98162722; PMID:9501991

A/Accession: T42516

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-427 <YOS>

A/Cross-references: EMBL:D89156; NID:g1749519; PIDN:BA13818.1; PID:g1749520

A/Experimental source: strain PR745

C/Superfamily: hypothetical protein SPBC17D11.01

Query Match Best Local Similarity 64.3%; Score 54; DB 2; Length 427;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
|:|:|:~|:|:~|:
Db 299 HHHHHHHHHHHH 310

RESULT 62

S66671

neuron-derived receptor NOR-2 - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 16-Aug-2004

C/Accession: S66671

R/Petropoulos, I.; Part, D.; Ochoa, A.; Zakari, M.M.; Lamas, E.

FEBS Lett. 372, 273-278, 1995

A/Title: NOR-2 (neuron-derived orphan receptor), a brain zinc finger protein, is highly

A/Reference number: S66671; MUID:96000221; PMID:755683

A/Accession: S66671

A/Status: translation not shown

A/Molecule type: mRNA

A/Residues: 1-430 <PET>

A/Cross-references: UNIPROT:P51179; EMBL:X86003; NID:g1065885; PIDN:CAA59993.1; PID:g10

A/Genetics: nor 2
C/Superfamily: erba transforming protein homology
C/Keywords: DNA binding; zinc finger
F/292-430/Domain: erba transforming protein homology #status atypical <ERBA>

F/292-379/Region: DNA binding

F/294-314/Region: zinc finger CCCC motif

F/312-316/Region: P box

F/330-354/Region: zinc finger CCCC motif

F/331-335/Region: D box

F/372-378/Region: A box

Query Match 64.3%; Score 54; DB 2; Length 430;
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S66148
 R:Medier, U.; Siegel, V.; Mlodzik, M.
 EMBL J. 14, 6247-6257, 1995
 A:Title: pipequeak encodes a novel nuclear protein required downstream of seven-up for t
 A:Reference number: S66148; MUID:96134923; PMID:8557044
 A:Accession: S66148
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-535 <WEB>
 A:Cross-references: UNIPROT:Q24455; EMBL:X90986; NID:g1149498; PIDN:CAA62473.1; PID:g1
 C:Genetics:
 F:21-123/Domain: po2 domain homology <PO2>

RESULT 63

C96704
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: C96704
 R:Theologis, A.; Ecker, J.R.; Palm, C.T.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 anson, N.F.; Hughes, B.; Hitzler, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: C96704
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-450 <STO>
 A:Cross-references: GB:AE005173; NID:g6553886; PIDN:AAAF16552.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: T23K23.5
 A:Map position: 1

Query Match 64.3%; Score 54; DB 2; Length 450;
 Best Local Similarity 50.0%; Pred. No. 1.3;
 Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 HNNHNNHNNH 12
 Db 47 HSHDHDHDDH 58

RESULT 64

S52453
 ecdysteroid UDP-glucosyltransferase - Spodoptera littoralis nuclear polyhedrosis virus
 C:Species: Spodoptera littoralis nuclear polyhedrosis virus, STMNV
 C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
 C:Accession: S52453
 R:Faktor, O.; Kamenskyy, B.
 submitted to the EMBL Data Library, February 1995
 A:Description: Nucleotide sequence of the ecdysteroid UDP-glucosyltransferase gene of Sp
 A:Reference number: S52453
 A:Accession: S52453
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-515 <PAK>
 A:Cross-references: UNIPROT:Q88168; EMBL:X84701; NID:g677862; PIDN:CAAS9174.1; PID:g6778

C:Superfamily: ecdysteroid UDPglucosyltransferase
 Query Match 64.3%; Score 54; DB 2; Length 515;
 Best Local Similarity 58.3%; Pred. No. 1.5;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Oy 1 HNNHNNHNNH 12
 Db 58 HNNHNNHNNH 69

Query Match 64.3%; Score 54; DB 2; Length 515;
 Best Local Similarity 58.3%; Pred. No. 1.5;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Oy 1 HNNHNNHNNH 12
 Db 58 HNNHNNHNNH 69

RESULT 65
 S66148
 gene pipequeak protein A short form - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S66148
 R:Medier, U.; Siegel, V.; Mlodzik, M.
 EMBL J. 14, 6247-6257, 1995
 A:Title: pipequeak encodes a novel nuclear protein required downstream of seven-up for t
 A:Reference number: S66148; MUID:96134923; PMID:8557044
 A:Accession: S66148
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-535 <WEB>
 A:Cross-references: UNIPROT:Q24455; EMBL:X90986; NID:g1149498; PIDN:CAA62473.1; PID:g1
 C:Genetics:
 F:21-123/Domain: po2 domain homology <PO2>

Query Match 64.3%; Score 54; DB 2; Length 535;
 Best Local Similarity 63.6%; Pred. No. 1.6;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Oy 1 HNNHNNHNNH 11
 Db 332 HEHNNHGHGH 342

RESULT 66

T49672
 related to a-agglutinin core protein AGA1 [imported] - Neurospora crassa
 N:Alternate names: protein B8B20.260
 C:Species: Neurospora crassa
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
 C:Accession: T49672
 R:Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatur
 submitted to the Protein Sequence Database, May 2000
 A:Reference number: Z25022
 A:Accession: T49672
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-590 <SCH>
 A:Cross-references: UNIPROT:Q9P5F9; EMBL:AL355933; GSPDB:GN00116; NCSP:B8B20.260
 A:Experimental source: BAC clone B8B20; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B8B20.260
 A:Map position: 6
 A:introns: 361/1

Query Match 64.3%; Score 54; DB 2; Length 590;
 Best Local Similarity 50.0%; Pred. No. 1.7;
 Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 HNNHNNHNNH 12
 Db 199 HSHSHSHSHS 210

RESULT 67

S71930
 neuron-derived receptor NOR-1 - human
 C:Species: Homo sapiens (man)
 C:Date: 04-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 20-Feb-1998
 C:Accession: S71930
 R:Ohkura, N.; Ito, M.; Tsukada, T.; Sasaki, K.; Yamaguchi, K.; Miki, K.
 Biochim. Biophys. Acta 1308, 205-214, 1996
 A:Title: Structure, mapping and expression of a human NOR-1 gene, the third member of th
 A:Reference number: S71930; MUID:96404972; PMID:8809112
 A:Accession: S71930
 A:Molecule type: mRNA
 A:Residues: 1-625 <OHK>
 A:Cross-references: EMBL:D78579
 A:Note: DNA was also sequenced
 C:Genetics:
 A:Map position: 9q
 C:Superfamily: probable hormone receptor NI0, nuclear; erba transforming protein homolog

Query Match 64.3%; Score 54; DB 2; Length 590;
 Best Local Similarity 50.0%; Pred. No. 1.7;
 Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 HNNHNNHNNH 12
 Db 199 HSHSHSHSHS 210

C/Keywords: DNA binding; receptor; zinc finger
F/290-544/Domain: erba transforming protein homology <ERBA>
F/290-377/Domain: DNA binding #status predicted <DNA>
F/292-312/Region: zinc finger CCCC motif
F/310-314/Region: P box
F/328-352/Region: zinc finger CCCC motif
F/329-333/Region: D box
F/370-376/Region: A box

Query Match 64.3%; Score 54; DB 2; Length 625;
Best Local Similarity 50.0%; Pred. No. 1.8;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:
DB 95 HHHHHHHHHH 106

RESULT 68

JC2493
neuron derived orphan receptor-1 - rat

N/Alternate names: NOR-1
C/Species: Rattus norvegicus (Norway rat)
C/Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C/Accession: JC2493; I48179

R/Okura, N.; Hijikuro, M.; Yamamoto, A.; Miki, K.

Biochem. Biophys. Res. Commun. 205, 1959-1965, 1994

A/Title: Molecular cloning of a novel thyroid/steroid receptor superfamily gene from cul

A/Reference number: JC2493; MUID:95110346; PMID:7811288

A/Accession: JC2493

A/Molecule type: mRNA

A/Residues: 1-628 <OHK>

A/Cross-references: UNIPROT:P51179; DDBJ:D38530; NID:G1483194; PID:BA07535.1; PID:g643

A/Experimental source: neuronal cell

R/Pena de Ortiz, S.; Cannon, M.M.; Jamieson, G.A.

Brain Res. Mol. Brain Res. 23, 278-283, 1994

A/Title: Expression of nuclear hormone receptors within the rat hippocampus: identificat

A/Reference number: I48178; MUID:94335560; PMID:7914660

A/Accession: I48179

A/Status: translated from GB/EMBL/DBD

A/Molecule type: mRNA

A/Residues: 318-345 <RES>

A/Cross-references: GB:I19343; NID:9450942; PID:AA646395.1; PID:g565379

A/Experimental source: hippocampus

C/Superfamily: probable hormone receptor N10, nuclear; erba transforming protein homolog

C/Keywords: DNA binding; receptor; zinc finger

F/292-547/Domain: erba transforming protein homology <ERBA>

F/292-379/Region: DNA binding #status predicted

F/312-316/Region: P box

F/331-335/Region: D box

F/372-378/Region: A box

Query Match 64.3%; Score 54; DB 2; Length 628;
Best Local Similarity 50.0%; Pred. No. 1.8;

Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:
DB 99 HHHHHHHHHH 110

RESULT 69

A26030
serine/threonine-specific protein kinase (EC 2.7.1.-) SNPL - yeast (Saccharomyces cerevi

N/Alternate names: protein YDR477w

C/Species: Saccharomyces cerevisiae

C/Date: 20-Aug-1987 #sequence_revision 20-Aug-1987 #text_change 09-Jul-2004

C/Accession: A26030; S69644

R/Celenza, J.L.; Carlson, M.

Science 233, 1175-1180, 1986

A/Title: A yeast gene that is essential for release from glucose repression encodes a pr

A/Reference number: A26030; MUID:86289463; PMID:3526554

A/Accession: A26030

A/Molecule type: DNA
A/Residues: 1-633 <CEL>
A/Cross-references: UNIPROT:P06782; EMBL:ML3971; NID:G172629; PID:AAA35058.1; PID:g172
R/Dietrich, F.S.
Submitted to the EMBL Data Library, August 1995
A/Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.

A/Reference number: S69554

A/Accession: S69644

A/Molecule type: DNA

A/Residues: 1-633 <DIE>

A/Cross-references: EMBL:U33050; NID:g927726; PID:AA64904.1; PID:g927732; GSPDB:GN000

C/Genetic:

A/Gene: SGD:SNF1; MIPS:YDR477w

A/Cross-references: SGD:S0002885; MIPS:YDR477w

A/Map position: 4R

C/Function:

A/Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threoni

A/Note: required for expression of glucose-repressed genes in response to glucose depriv

C/Superfamily: AMP-activated protein kinase; protein kinase homology

C/Keywords: ATP; autophosphorylation; magnesium; nucleus; phosphoprotein; phosphotransf

F/53-106/Domain: protein kinase homology <KIN>

F/61-69/Region: protein kinase ATP-binding motif

F/84,103,177,179/Active site: Lys, Glu, Asp, Lys #status predicted

F/182,186/Binding site: magnesium (Zn, Asp) #status predicted

F/210/Binding site: phosphate (Thr) (covalent) (by unidentified kinase) #status predict

Query Match 64.3%; Score 54; DB 1; Length 633;
Best Local Similarity 50.0%; Pred. No. 1.8;

Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:
DB 18 HHHHHHHHHH 29

RESULT 70

JC4898
Down-syndrome-critical-region protein - human

N/Alternate names: Drosophila minibrain protein homolog

C/Species: Homo sapiens (man)

C/Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 16-Aug-2004

C/Accession: JC4898

R/Shindoh, N.; Kudoh, J.; Maeda, H.; Yamaki, A.; Minoshima, S.; Shimizu, Y.; Shimizu, N

Biochem. Biophys. Res. Commun. 225, 92-99, 1996

A/Title: Cloning of a human homolog of the Drosophila minibrain/rat dyrk gene from "the

A/Reference number: JC4898; MUID:96332410; PMID:8769099

A/Accession: JC4898

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-754 <SHI>

A/Cross-references: UNIPROT:Q13627; DDBJ:DB5759; NID:G1526445; PID:BAAL2866.1; PID:g15

C/Superfamily: protein kinase homology

Query Match 64.3%; Score 54; DB 2; Length 754;
Best Local Similarity 50.0%; Pred. No. 2.2;

Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:
DB 598 HHHHHHHHHH 609

RESULT 71

AH3258
cation-transporting ATPase pacs (EC 3.6.1.-) [imported] - Brucella melitensis (strain 16

C/Species: Brucella melitensis

C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C/Accession: AH3258

R/Delvecchio, V.G.; Kapatal, V.; Redkar, R.V.; Patra, G.; Muller, C.; Los, T.; Ivanova,

.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Leteas

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A/Title: The genome sequence of the facultative intracellular pathogen Brucella melitens

A/Reference number: AH3252; PMID:11756688

Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNHN 12
|:|:|:|:|:
Db 598 HHHHHHHHHH 609

RESULT 76

C64698

probable histidine-rich metal-binding protein - *Helicobacter pylori*

C/Species: *Helicobacter pylori*

A/Variety: strains J99, 26695

C/Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004

C/Accession: C64698; C71821

R/Tomb: J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, L.

Nature 388, 539-547, 1997

A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.

A/Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.

A/Reference number: A64520; MUID:97394467; PMID:9252185

A/Accession: C64698

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-60 <TOM>

A/Cross-references: UNIPROT:Q48251; GB:AE000643; GB:AE000511; NID:G2314598; PIDN:AAD0847

C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 12-Jul-2004

C/Accession: G64075

R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.;

Gokeyne, J.D.; Scott, J.; Shirley, R.; Liu, L.; Glodek, A.; Kelley, J.M.; Weidman, J.D.M.;

Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M.

Science 269, 496-512, 1995

A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.

A/Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.

A/Reference number: A64000; MUID:95350630; PMID:7542800

A/Accession: G64075

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-185 <TRIG>

A/Cross-references: UNIPROT:P44394; GB:U32736; GB:L42023; NID:G1573519; PIDN:AC22196.1

C/Suprafamily: urease accessory protein (nickel metallochaperone) Ure

A/KeyWords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulation

QY 1 HNNNNNNNNHN 12
|:|:|:|:|:
Db 11 HHHHHHHHHH 22

Query Match 63.1%; Score 53; DB 2; Length 60;

Best Local Similarity 50.0%; Pred. No. 0.23;

Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNNNNNNNHN 12
|:|:|:|:|:
Db 11 HHHHHHHHHH 22

RESULT 77

B30242

stem cell protein ERA-1-399, retinoic acid-induced - mouse

C/Species: *Mus musculus* (house mouse)

C/Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004

C/Accession: B30242

R/Larosa, G.V.; Gudas, L.J.

Mol. Cell. Biol. 8, 3906-3917, 1988

A/Title: Early retinoic acid-induced P9 teratocarcinoma stem cell gene ERA-1: alternate

A/Reference number: A30242; MUID:9127233; PMID:2906112

A/Accession: B30242

A/Molecule type: mRNA

A/Residues: 1-133 <LAR>

A/Cross-references: UNIPROT:P09022; GB:M22115; NID:G13047; PIDN:AAA37558.1; PID:G387146

C/Suprafamily: unassigned homeobox proteins; homeobox homology

A/KeyWords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulation

Query Match 63.1%; Score 53; DB 2; Length 133;

Best Local Similarity 54.5%; Pred. No. 0.52;

Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNH 11
|:|:|:|:|:
Db 60 HHHHHHHHHH 70

RESULT 78

B64421

conserved hypothetical protein M0970 - *Methanococcus jannaschii*

C/Species: *Methanococcus jannaschii*

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 12-Jul-2004

C/Accession: B64421

R/Bull, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, A.

Science 273, 1058-1073, 1996

A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.

A/Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*

A/Reference number: A64300; MUID:96337999; PMID:8688087

A/Accession: B64421

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-143 <BUL>

A/Cross-references: UNIPROT:Q58380; GB:U67540; GB:L77117; NID:G1591631; PIDN:AA98975.1

C/Suprafamily: cobaltochelatase/ferrochelatase CdxX/SirB

A/KeyWords: cobaltochelatase/ferrochelatase CdxX/SirB

QY 1 HNNNNNNNNHN 12
|:|:|:|:|:
Db 93 HHHHHHHHHH 104

Query Match 63.1%; Score 53; DB 1; Length 143;

Best Local Similarity 50.0%; Pred. No. 0.55;

Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNNNNNNNHN 12
|:|:|:|:|:
Db 93 HHHHHHHHHH 104

Query Match 63.1%; Score 53; DB 2; Length 185;

Best Local Similarity 50.0%; Pred. No. 0.72;

Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNNNNNNNHN 12
|:|:|:|:|:
Db 161 HGHSHSHSHDH 172

RESULT 80

S74993

hypothetical protein s11011 - *Synechocystis* sp. (strain PCC 6803)

C/Species: *Synechocystis* sp.

C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C/Accession: S74993

R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yanada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

5
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74993
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-270 <KAN>
A:Cross-references: UNIPROT:P73013; EMBL:D90902; GB:AB001339; NID:g1652027; PIDD:BA11703
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Start codon: GTG
F:14-84/Domain: dnaJ amino-terminal homology <DNJ>

Query Match 63.1%; Score 53; DB 2; Length 270;
Best Local Similarity 58.3%; Pred. No. 1.1;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:
DB 123 HHHHHQTHNNHS 134

RESULT 81
H90992
hypothetical protein ECG2912 [imported] - Escherichia coli (strain O157:H7, substrain R1
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: H90992
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A93629; MUID:21156231; PMID:11258796
A:Accession: H90992
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-279 <HAY>
A:Cross-references: UNIPROT:Q8X369; GB:BA000007; PIDD:BA36335.1; PID:g13362381; GSPDB:G
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECG2912

Query Match 63.1%; Score 53; DB 2; Length 279;
Best Local Similarity 50.0%; Pred. No. 1.1;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:
DB 134 HDHDDHDEHHNN 145

RESULT 82
C85838
hypothetical protein yohM [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: C85838
R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaeser, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: C85838
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-283 <STO>
A:Cross-references: UNIPROT:Q8X3U5; GB:AE005174; NID:g12516319; PIDD:AAG57167.1; GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yohM

Query Match 63.1%; Score 53; DB 2; Length 283;
Best Local Similarity 54.5%; Pred. No. 1.1;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 11
|:|:|:|:|:|:
DB 134 HDHDDHDDHDDH 144

RESULT 83
A55641
homeotic protein GAX - human
C:Species: Homo sapiens (man)
C>Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 09-Jul-2004
C:Accession: A55641
R:LePage, D.F.; Allomare, D.A.; Testa, J.R.; Walsh, K.
Genomics 24, 535-540, 1994
A:Title: Molecular cloning and localization of the human GAX gene to 7p21.
A:Reference number: A55641; MUID:95229154; PMID:7713505
A:Accession: A55641
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-302 <LEP>
A:Cross-references: UNIPROT:P50222; GB:I36328; NID:g557549; PIDD:AA58497.1; PID:g55755
C:Genetics:
A:Gene: GAX
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:186-242/Domain: homeobox homology <HOX>

Query Match 63.1%; Score 53; DB 2; Length 302;
Best Local Similarity 54.5%; Pred. No. 1.2;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 11
|:|:|:|:|:|:
DB 68 HHHHHHHHHNNH 78

RESULT 84
A56235
transcription activator MafB - chicken
C:Species: Gallus gallus (chicken)
C>Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C:Accession: A56235
R:Katohka, K.; Fujiwara, K.T.; Noda, M.; Nishizawa, M.
Mol. Cell. Biol. 14, 7581-7591, 1994
A:Title: MafB, a new Maf family transcription activator that can associate with Maf and
A:Reference number: A56235; MUID:95021288; PMID:7935473
A:Accession: A56235
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-311 <KAT>
A:Cross-references: UNIPROT:Q90888; GB:I28600; NID:g516723; PIDD:BA05938.1; PID:g516724
C:Genetics:
A:Introns: #status absent
C:Superfamily: maf transforming protein; maf homology
C:Keywords: DNA binding; homeodimer; leucine zipper
F:200-289/Domain: maf homology <MAF>

Query Match 63.1%; Score 53; DB 2; Length 311;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:
DB 130 HHHHHHHHHNNH 141

RESULT 85
A30242
homeotic protein ERA-1-993, retinoic acid-induced - mouse
C:Species: Mus musculus (house mouse)
C>Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C:Accession: A30242
R:Ikarova, G.J.; Gudae, L.J.

Mol. Cell. Biol. 8, 3906-3917, 1988
A>Title: Early retinoid acid-induced P9 teratocarcinoma stem cell gene ERA-1: alternate
A;Reference number: A30242; MUID:89127233; PMID:2906112
A;Accession: A30242
A;Molecule type: mRNA
A;Residues: 1-331 <LAR>
A;Cross-references: UNIPROT:P09022; GB:M22115; NID:G193047; PIDN:AAA37559.1; PID:G3092218
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulat
F;226-282/Domain: homeobox homology <HOX>

Query Match 63.1%; Score 53; DB 2; Length 331;
Best Local Similarity 54.5%; Pred. No. 1.3;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
|:|:|:|:|:|:
DB 60 HHHHHHHHHH 70

RESULT 86
D64049
adhesin homolog HI0119 - Haemophilus influenzae (strain Rd KW20)

C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C;Accession: D64049
R;Fleischmann, D.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kiehlavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodet, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: D64049

A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-337 <TIGR>

A;Cross-references: UNIPROT:P44526; GB:U32698; GB:U42023; NID:G3212178; PIDN:AAC21794.1;
C;Superfamily: hypothetical protein HI0119

Query Match 63.1%; Score 53; DB 2; Length 337;
Best Local Similarity 50.0%; Pred. No. 1.3;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:|:
DB 135 HDKHDHDDHD 146

RESULT 87

E83340
hypothetical protein PA2438 [imported] - Pseudomonas aeruginosa (strain PAO1)

C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: E83340

R;Stover, C.K.; Pham, X.Q.; Errin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83340

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-341 <STO>

A;Cross-references: UNIPROT:Q91144; GB:AE004671; GB:AE004091; NID:G9948483; PIDN:AAG0562
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA2438

Query Match 63.1%; Score 53; DB 2; Length 341;
Best Local Similarity 50.0%; Pred. No. 1.3;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:|:
DB 12 HHHHHDDHDDHD 23

RESULT 88

KG20HL
histidine-rich glycoprotein precursor - Plasmodium lophurae

C;Species: Plasmodium lophurae
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jul-2004
C;Accession: A22692

R;Ravech, J.V.; Feder, R.; Pavlovic, A.; Blobel, G.
Nature 312, 616-620, 1984
A;Title: Primary structure and genomic organization of the histidine-rich protein of th

A;Reference number: A22692; MUID:85061618; PMID:6095114

A;Accession: A22692

A;Molecule type: DNA
A;Residues: 1-351 <RAV>

A;Cross-references: UNIPROT:P04929; GB:X01469; NID:G9997; PIDN:CAA25698.1; PID:G9999
C;Comment: There are two copies of 16-residue repeats, two copies of 17-residue repeats
C;Genetics:

A;Introns: 23/3

C;Superfamily: plasmodium histidine-rich protein

C;Keywords: glycoprotein; tandem repeat
F;1-23/Domain: signal sequence #status predicted <SIG>

F;24-47/Domain: propeptide #status predicted <PRO>

F;48-351/Product: histidine-rich glycoprotein #status predicted <MAT>

F;53-74/Region: 16-residue repeats
F;91-107/Region: 17-residue repeats

F;124-138/Region: 15-residue repeats
F;173-301/Region: 10-residue repeats

F;40/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 63.1%; Score 53; DB 1; Length 351;
Best Local Similarity 50.0%; Pred. No. 1.4;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:|:
DB 216 HHHHHHHHHH 227

RESULT 89

AB2396
hypothetical protein a114722 [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AB2396

R;Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuch

Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, T.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AB2396

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-351 <RUR>

A;Cross-references: UNIPROT:Q8YNA8; GB:BA000019; PIDN:BAH76421.1; PID:G17133859; GSPDB:

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: a114722

C;Superfamily: conserved hypothetical protein ylc1

Query Match 63.1%; Score 53; DB 2; Length 353;
Best Local Similarity 50.0%; Pred. No. 1.4;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:|:
DB 235 HSHHDDHDDHD 246

RESULT 90

S35345
C:Species: Mus musculus (house mouse)
C:Date: 08-Jun-1994 #sequence_revision 10-Nov-1995 #text_change 16-Aug-2004
C:Accession: S35345; S20087
R:Simone, A.; Acampora, D.; Mallamaci, A.; Stornaiuolo, A.; d'Apice, M.R.; Nigro, V.; B
EMBO J. 12, 2735-2747, 1993
A:Title: A vertebrate gene related to orthodenticle contains a homeodomain of the bicoid
A:Reference number: S35345; MUID:93327763; PMID:8101484
A:Accession: S35345
A:Molecule type: DNA
A:Residues: 1-355 <SIM>
A:Cross-references: UNIPROT:P80205
R:Simone, A.; Acampora, D.; Gulsano, M.; Stornaiuolo, A.; Boncinelli, E.
Nature 358, 687-690, 1992
A:Title: Nested expression domains of four homeobox genes in developing rostral brain.
A:Reference number: S29084; MUID:92357147; PMID:1353865
A:Accession: S29087
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 31-98 <S12>
A:Cross-references: EMBL:X68883; NID:953540; PIDN:CAA46754.1; PID:953541
C:Gene: cdx1
C:Superfamily: homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:39-95/Domain: homeobox homology <HOX>

Query Match

Best Local Similarity 63.1%; Score 53; DB 2; Length 355;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHHNNHHNNH 11
|:|:|:|:|:
Db 292 HHHHHHHHHH 302

RESULT 91

156547
homeodomain protein cdx1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 16-Aug-2004
C:Accession: 156547
R:Francz, G.D.; Weimann, J.M.; Levin, M.E.; McConnell, S.K.
J. Neurosci. 14, 5725-5740, 1994
A:Title: Cdx1 and Cdx2 define layers and regions in developing cerebral cortex and cereb
A:Reference number: 156547; MUID:95016961; PMID:7931541
A:Accession: 156547
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-355 <RES>
A:Cross-references: UNIPROT:Q63410; GB:L32602; NID:9535739; PIDN:AAAS3557.1; PID:9535740
C:Gene: OTX1
C:Superfamily: homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:39-95/Domain: homeobox homology <HOX>

Query Match
Best Local Similarity 63.1%; Score 53; DB 2; Length 355;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHHNNHHNNH 11
|:|:|:|:|:
Db 292 HHHHHHHHHH 302

RESULT 92

S76964
hypothetical protein - Synecchocystis sp. (strain PCC 6803)
C:Species: Synecchocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S76964
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N
o, K.; Okunura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yaeud
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecchocysti
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76964
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-383 <KAN>
A:Cross-references: UNIPROT:P74756; EMBL:D90917; GB:AB001339; NID:91653836; PIDN:BAA188
C:Superfamily: conserved hypothetical protein ytcC

Query Match

Best Local Similarity 63.1%; Score 53; DB 2; Length 383;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 HHHNNHHNNH 12
|:|:|:|:|:
Db 228 HHHSHDHGHDH 239

RESULT 93

A40350
transcription repressor protein YY1 - human
N:Alternate names: transcription repressor/activator protein NF-E1
C:Species: Homo sapiens (man)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C:Accession: A40350; S78494; S33712; A56419
R:Shi, Y.; Seto, E.; Chang, L.S.; Shenk, T.
Cell 67, 377-388, 1991
A:Title: Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and re
A:Reference number: A40350; MUID:92005726; PMID:1655281
A:Accession: A40350
A:Molecule type: mRNA
A:Residues: 1-414 <SHI>
A:Cross-references: UNIPROT:P25490; GB:M77698; NID:9186767; PIDN:AAAS9467.1; PID:9186768
A:Experimental source: HeLa cells
A>Note: The authors translated the codon CGC for residue 371 as Lys, CCA for residue 375
submitted to the EMBL Data Library, July 1992.
A:Description: Observed and predicted DNA binding of a zinc finger protein which recogni
A:Reference number: S78494
A:Accession: S78494

A:Molecule type: mRNA
A:Residues: 1-195, 'G', 197-414 <WHI>
A:Cross-references: EMBL:Z14077; NID:938010; PIDN:CAA78455.1; PID:938011
R:Park, K.; Atchison, M.L.
Proc. Natl. Acad. Sci. U.S.A. 89, 9804-9808, 1991
A:Title: Isolation of a candidate repressor/activator, NF-E1 (YY-1), that binds to
A:Reference number: A56419; MUID:92052179; PMID:1946405
A:Accession: S33712
A:Molecule type: mRNA
A:Residues: 1-64, 'R', 66-195, 'G', 197-414 <PAR>
A:Cross-references: GB:M76541; NID:9189173; PIDN:AAAS9926.1; PID:9189174
C:Keywords: DNA binding; transcription regulation; zinc finger
F:298-320/Region: zinc finger CCH motif
F:327-347/Region: zinc finger CCH motif
F:355-377/Region: zinc finger CCH motif
F:385-407/Region: zinc finger CCH motif

Query Match
Best Local Similarity 63.1%; Score 53; DB 2; Length 414;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHHNNHHNNH 11
|:|:|:|:|:
Db 70 HHHHHHHHHH 80

RESULT 94
C96995
Uncharacterized conserved protein CAC0774 [imported] - Clostridium acetobutylicum
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C/Accession: C96995
R/Molling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium acetobutylicum*
A/Reference number: A96900; MUID:21359325; PMID:21359325
A/Accession: C96995
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-420 <KUR>
A/Cross-references: UNIPROT:Q97KZ2; GB:AE001437; PIDD:AAK78750.1; PID:G15023659; GSPDB:G
A/Experimental source: Clostridium acetobutylicum ATCC824
C/Genetic8:
A/Genetic: CAC0774
C/Superfamily: Methanobacterium thermoautotrophicum hypothetical protein MTH1109

Query Match 63.1%; Score 53; DB 2; Length 420;
Best Local Similarity 70.0%; Pred. No. 1.7;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 10
| | | | |
| | | | |
Db 80 HNNHNNHNNH 89

RESULT 95
I37451
HBF-G2 (HFK-2) protein - human
C/Species: Homo sapiens (man)
C/Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 16-Aug-2004
C/Accession: I37451
R/Wiese, S.; Murphy, D.B.; Schlund, A.; Burfeind, P.; Schmundt, D.; Schunle, V.; Mattei
Biochim. Biophys. Acta 1262, 105-112, 1995
A/Title: The genes for human brain factor 1 and 2, members of the fork head gene family,
A/Reference number: I37451; MUID:95322450; PMID:7599184
A/Accession: I37451
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-469 <RES>
A/Cross-references: UNIPROT:P55316; EMBL:X78202; NID:9967047; PIDD:CAAS5038.1; PID:99670
C/Genetic8:
A/Genetic: HBF-G2; HFK-2
A/Suprafamily: fork head DNA-binding domain homology
F:162-253/Domain: fork head DNA-binding domain homology <FHD>

Query Match 63.1%; Score 53; DB 2; Length 469;
Best Local Similarity 54.5%; Pred. No. 1.8;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
| | | | |
| | | | |
Db 47 HNNHNNHNNH 57

RESULT 96
A54743
transcription factor HFK1 - human
C/Species: Homo sapiens (man)
C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 12-Sep-1997
C/Accession: A54743
R/Murphy, D.B.; Wiese, S.; Burfeind, P.; Schmundt, D.; Mattei, M.G.; Schulz-Schaeffer, W
Genomics 21, 551-557, 1994
A/Title: Human brain factor 1, a new member of the fork head gene family.
A/Reference number: A54743; MUID:95048332; PMID:7959731
A/Accession: A54743
A/Status: preliminary
A/Molecule type: mRNA

A/Residues: 1-476 <MUR>
A/Cross-references: GB:X74142
C/Genetic8:
A/Genetic: GDB:FKH14; HBF-1; HFK1
A/Cross-references: GDB:433550
A/Map position: 14q12-14q12
C/Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
F:169-260/Domain: fork head DNA-binding domain homology <FHD>

Query Match 63.1%; Score 53; DB 2; Length 476;
Best Local Similarity 54.5%; Pred. No. 1.9;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
| | | | |
| | | | |
Db 47 HNNHNNHNNH 57

RESULT 97
S71342
calnexin precursor - Korean frog
C/Species: Rana rugosa (Korean frog)
C/Date: 25-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 09-Jul-2004
C/Accession: S71342
R/Yamamoto, S.; Nakamura, M.
FEBS Lett. 387, 27-32, 1996
A/Title: Calnexin: its molecular cloning and expression in the liver of the frog, *Rana*
A/Reference number: S71342; MUID:96234004; PMID:8654561
A/Accession: S71342
A/Molecule type: mRNA
A/Residues: 1-622 <YAM>
A/Cross-references: UNIPROT:Q98985; EMBL:D78590; NID:G1514958; PIDD:BA11426.1; PID:G15
A/Experimental source: liver
C/Superfamily: calnexin
C/Keywords: calcium binding; endoplasmic reticulum; molecular chaperone; transmembrane
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-622/Product: calnexin #status predicted <MAT>
F:478-500/Domain: transmembrane #status predicted <TMW>

Query Match 63.1%; Score 53; DB 2; Length 622;
Best Local Similarity 50.0%; Pred. No. 2.5;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
| | | | |
| | | | |
Db 26 HNNHNNHNNH 37

RESULT 98
S33167
gene pointed protein splice form 1 - fruit fly (*Drosophila melanogaster*)
N/Alternate names: FTS-like protein
C/Species: *Drosophila melanogaster*
C/Date: 13-Jan-1995 #sequence_revision 30-May-1997 #text_change 09-Jul-2004
C/Accession: S33167; S28819; S28818
R/Klaembt, C.
Development 117, 163-176, 1993
A/Title: The *Drosophila* gene pointed encodes two FTS-like proteins which are involved in
A/Reference number: S33167; MUID:94038653; PMID:8223245
A/Accession: S33167
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-623 <KLA>
A/Cross-references: UNIPROT:P51022; EMBL:X69166; NID:G288079; PIDD:CAA48916.1; PID:G2880
R/Chen, T.; Bunting, M.; Karim, F.D.; Thummel, C.S.
Dev. Biol. 151, 176-191, 1992
A/Title: Isolation and characterization of five *Drosophila* genes that encode an ets-rel
A/Reference number: S28819; MUID:92249640; PMID:1577186
A/Accession: S28819
A/Molecule type: mRNA
A/Residues: 456-613 <CHE>
A/Cross-references: EMBL:M88472; NID:G157191; PIDD:AA34200.1; PID:G157192
R/Frithy, L.J.; Watson, D.K.; McWilliam, M.J.; Ascione, R.; Papae, T.S.

Search completed: October 4, 2005, 11:47:58
Job time : 53 secs

Dev. Biol. 127, 45-53, 1988

A:Title: The *Drosophila* ets-2 gene: molecular structure, chromosomal localization, and d

A:Reference number: S28818; MUID:88196618; PMID:2834248

A:Accession: S28818

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 445-466, 'R', 468-603 <PRI>

A:Cross-references: EMBL:M20408; NID:g157384; PIDN:AA28521.1; PID:g157385

A>Note: the authors translated the codon AAG for residue 467 as Gln

C:Genetics:

A:Gene: FLYBase:pnt

A:Cross-references: FLYBase:FBgn0003118

A:Introns: 550/3

C:Superfamily: pointed protein, splice form 1; ets DNA-binding domain homology
C:Keywords: alternative splicing; DNA binding; nucleus; transcription factor
F:517-595/Domain: ets DNA-binding domain homology <ETS>

Query Match

Best Local Similarity 63.1%; Score 53; DB 1; Length 623;

Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11

||:|:|:|:|:|

Db 75 HHHHHHHHSH 85

RESULT 99

A29454

knob-associated histidine-rich protein precursor - malaria parasite (*Plasmodium falcipar*

C:Species: *Plasmodium falciparum*

C>Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004

C:Accession: A29454

R:Trigila, T.; Stahl, H.D.; Crewther, P.E.; Scanlon, D.; Brown, G.V.; Anders, R.F.; Kemp

EMBO J. 6, 1413-1419, 1987

A:Title: The complete sequence of the gene for the knob-associated histidine-rich protei

A:Reference number: A29454; MUID:8725836; PMID:3301325

A:Accession: A29454

A:Molecule type: DNA

A:Residues: 1-657 <TRI>

A:Cross-references: UNIPROT:P06719; GB:Y00060; NID:g9908; PIDN:CAA68268.1; PID:g9909

C:Superfamily: knob-associated histidine-rich protein

Query Match

Best Local Similarity 63.1%; Score 53; DB 2; Length 657;

Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12

||:|:|:|:|:|

Db 61 HHHHHHHHHH 72

RESULT 100

S14535

asparagine-rich protein (clone 28C4) - malaria parasite (*Plasmodium falciparum*)

C:Species: *Plasmodium falciparum*

C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004

C:Accession: S14535

R:Schreibler, L.; Deutsche, U.; Storch, T.; Mueller-Hill, D.

Submitted to the EMBL Data Library, December 1989

A:Reference number: S14469

A:Accession: S14535

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-669 <SCH>

A:Cross-references: UNIPROT:Q25771; EMBL:X17486; NID:g9848; PID:g9849

Query Match

Best Local Similarity 63.1%; Score 53; DB 2; Length 669;

Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12

||:|:|:|:|:|

Db 523 HHHHHNNNNN 534

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2005, 11:22:37 ; Search time 120 Seconds
(without alignments)
51.208 Million cell updates/sec

Title: US-09-858-332G-15
Perfect score: 84
Sequence: 1 HNHNNHNNHNNH 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Uniprot_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	100.0	120	2 Q6JIT2	Q6JIT2 rattus norv
2	84	100.0	130	2 Q6PD05	Q6PD05 rattus norv
3	84	100.0	142	2 Q8T8F1	Q8T8F1 plasmodium
4	84	100.0	362	2 Q6Z1Z3	Q6Z1Z3 oryza sativ
5	84	100.0	362	2 Q7VAF5	Q7VAF5 prochloroco
6	84	100.0	439	2 Q9HEX9	Q9HEX9 pneumocysti
7	84	100.0	441	2 Q6CIY7	Q6CIY7 yarrowia li
8	84	100.0	1125	2 Q8G1F5	Q8G1F5 dictyosteli
9	84	100.0	2072	2 Q7RSW8	Q7RSW8 plasmodium
10	84	100.0	2969	2 Q7KFT3	Q7KFT3 plasmodium
11	84	100.0	2976	2 Q8WRS4	Q8WRS4 plasmodium
12	84	100.0	2977	2 Q8WPS9	Q8WPS9 plasmodium
13	78	92.9	417	2 Q7S0U3	Q7S0U3 neurospora
14	78	92.9	834	1 MPT5_YEAST	P39016 saccharomyc
15	76	90.5	102	1 YAC1_MAIZE	P08771 zea mays (m
16	76	90.5	501	2 Q815I6	Q815I6 plasmodium
17	74	88.1	747	1 NEUR_DROVI	Q24746 drosophila
18	74	88.1	1544	2 Q8MQWO	Q8MQWO drosophila
19	73.5	87.5	800	2 Q43988	Q43988 dictyosteli
20	73.5	87.5	1670	2 Q23901	Q23901 dictyosteli
21	73.5	87.5	1671	2 Q865R5	Q865R5 dictyosteli
22	71	84.5	1210	2 Q815M3	Q815M3 plasmodium
23	71	84.5	1210	2 Q7ROA5	Q7ROA5 plasmodium
24	70	83.3	233	2 Q9P826	Q9P826 candida alb
25	70	83.3	490	1 CAR3_DICDI	P35352 dictyosteli
26	70	83.3	942	2 Q869W0	Q869W0 dictyosteli
27	69	82.1	72	2 Q8T1A1	Q8T1A1 dictyosteli
28	69	82.1	1365	2 Q9VJ79	Q9VJ79 drosophila
29	69	82.1	2770	2 Q8T1V0	Q8T1V0 plasmodium
30	68.5	81.5	753	2 Q9VZX2	Q9VZX2 drosophila
31	68	81.0	136	2 Q8T8F2	Q8T8F2 plasmodium

32	68	81.0	358	2 Q6VZ16	Q6VZ16 canarypox v
33	68	81.0	584	2 Q869Z3	Q869Z3 dictyosteli
34	68	81.0	604	2 Q7S9S3	Q7S9S3 neurospora
35	68	81.0	2965	2 Q8WRS5	Q8WRS5 plasmodium
36	67	79.8	89	2 Q6A1P7	Q6A1P7 euplotes va
37	67	79.8	105	2 Q8DWM2	Q8DWM2 streptococ
38	67	79.8	105	2 Q8B2S4	Q8B2S4 streptococ
39	67	79.8	132	2 Q8T8F3	Q8T8F3 plasmodium
40	67	79.8	821	2 Q6C3S6	Q6C3S6 yarrowia li
41	67	79.8	898	2 Q7RUV4	Q7RUV4 neurospora
42	67	79.8	1107	2 Q7RH05	Q7RH05 plasmodium
43	67	79.8	2357	2 Q8WRS6	Q8WRS6 plasmodium
44	67	79.8	114	2 Q9D1N5	Q9D1N5 mus musculi
45	65	77.4	124	2 Q9CMN9	Q9CMN9 mus musculi
46	65	77.4	124	2 Q9WTZ9	Q9WTZ9 mus musculi
47	65	77.4	124	2 Q66GY9	Q66GY9 xenopus lae
48	65	77.4	147	2 Q9D0S2	Q9D0S2 mus musculi
49	65	77.4	303	2 Q8Y4B0	Q8Y4B0 listeria mo
50	65	77.4	303	2 Q927R8	Q927R8 listeria in
51	65	77.4	303	2 Q71WK3	Q71WK3 listeria mo
52	65	77.4	374	2 Q6CKC9	Q6CKC9 kluyveromyc
53	65	77.4	1010	2 Q9U0J3	Q9U0J3 plasmodium
54	65	77.4	1283	2 Q8T1U7	Q8T1U7 plasmodium
55	64	76.2	244	2 Q9S1B4	Q9S1B4 arabidopsis
56	64	76.2	308	2 Q72M12	Q72M12 leptospira
57	64	76.2	308	2 Q8E248	Q8E248 leptospira
58	64	76.2	329	2 Q945N2	Q945N2 arabidopsis
59	64	76.2	451	2 Q51561	Q51561 borrelia bu
60	64	76.2	545	2 Q80567	Q80567 arabidopsis
61	64	76.2	576	2 Q81D23	Q81D23 plasmodium
62	64	76.2	684	2 Q9SG87	Q9SG87 arabidopsis
63	64	76.2	786	2 Q86A52	Q86A52 dictyosteli
64	64	76.2	825	2 Q81KM1	Q81KM1 plasmodium
65	64	76.2	1175	1 HMDH_PHYBL	Q12649 phycomyc
66	64	76.2	1185	2 Q7KMF2	Q7KMF2 dictyosteli
67	64	76.2	1709	2 Q95PH5	Q95PH5 dictyosteli
68	63	75.0	118	2 Q84N72	Q84N72 amicia glan
69	63	75.0	250	2 Q6AN47	Q6AN47 desulfotale
70	63	75.0	341	1 ACER2_TIRRE	Q66wn6 trichoderma
71	63	75.0	341	2 Q81MS9	Q81MS9 drosophila
72	63	75.0	380	2 Q7S7N8	Q7S7N8 neurospora
73	63	75.0	400	2 Q871H5	Q871H5 neurospora
74	63	75.0	592	2 Q6LFF6	Q6LFF6 plasmodium
75	63	75.0	753	2 Q811O8	Q811O8 plasmodium
76	63	75.0	941	2 Q865Y0	Q865Y0 dictyosteli
77	63	75.0	1088	2 Q815S4	Q815S4 plasmodium
78	63	75.0	1097	2 Q7S316	Q7S316 neurospora
79	63	75.0	1348	2 Q81XK9	Q81XK9 plasmodium
80	63	75.0	1425	2 Q81X08	Q81X08 plasmodium
81	63	75.0	1621	2 Q81LL7	Q81LL7 plasmodium
82	63	75.0	1840	2 Q81ED3	Q81ED3 plasmodium
83	63	75.0	2075	2 Q81HRS	Q81HRS plasmodium
84	63	75.0	2461	2 Q6LFF7	Q6LFF7 plasmodium
85	63	75.0	2506	2 Q812Y8	Q812Y8 plasmodium
86	63	75.0	2558	2 Q815J1	Q815J1 plasmodium
87	63	75.0	5561	2 Q81BHS	Q81BHS plasmodium
88	63	75.0	5890	2 Q81K84	Q81K84 plasmodium
89	63	75.8	329	2 Q9M2K4	Q9M2K4 arabidopsis
90	62	73.8	529	2 Q9Y3Z1	Q9Y3Z1 homo sapien
91	62	73.8	601	1 KEN_DROME	Q9Y3Z1 homo sapien
92	62	73.8	732	1 YLH3_SCHPO	Q9Y3Z1 schizosach
93	62	73.8	747	2 Q7RML2	Q7RML2 plasmodium
94	62	73.8	831	2 Q68CR5	Q68CR5 homo sapien
95	62	73.8	835	2 Q9ULP5	Q9ULP5 homo sapien
96	62	73.8	911	2 Q86H89	Q86H89 dictyosteli
97	62	73.8	1342	2 Q81218	Q81218 plasmodium
98	61	72.6	190	1 SLVD_HAEIN	P44830 haemophilus
99	61	72.6	271	2 Q94451	Q94451 dictyosteli
100	61	72.6	607	2 Q7RCN8	Q7RCN8 plasmodium

ALIGNMENTS

```

RESULT 1
O9JIT2
ID O9JIT2 PRELIMINARY; PRT; 120 AA.
AC O9JIT2.
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE p75NTR-associated cell death executor.
GN Name=Nade;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA MEDLINE=20298829; PubMed=10764727; DOI=10.1074/jbc.C000140200;
RA Mukai J., Hachiya T., Shoji-Hoshino S., Kimura M., Nadano D.,
RA Suvarito P., Hanaoka T., Li Y., Irie S., Greene L.A., Sato T.A.;
RT "NADP, a p75NTR-associated cell death executor, is involved in signal
RT transduction mediated by the common neurotrophin receptor p75NTR."
RL J. Biol. Chem. 275:17566-17570(2000).
DR EMBL; AF187065; AAF5130.1; -.
DR InterPro; IPR007623; BEX.
DR Pfam; PF04538; BEX; 1.
SQ SEQUENCE 120 AA; 14137 MW; EF7541D3DB6638B CRC64;

Query Match
Best Local Similarity 100.0%; Score 84; DB 2; Length 120;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 12
Db 29 HNNHNNHNNHNN 40

RESULT 2
O6PDUS
ID O6PDUS PRELIMINARY; PRT; 130 AA.
AC O6PDUS.
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Ngr1ap1 protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary gland;
RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.U., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buttefield Y.S.,
RA Kozajski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Maira M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary gland;
RA Strausberg R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC058503; AAH58503.1; -.
SQ SEQUENCE 130 AA; 15333 MW; 578BD9CF343E934A CRC64;

Query Match
Best Local Similarity 100.0%; Score 84; DB 2; Length 130;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 12
Db 39 HNNHNNHNNHNN 50

RESULT 3
O8T8F1
ID O8T8F1 PRELIMINARY; PRT; 142 AA.
AC O8T8F1.
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Rhl1 protein (fragment).
GN Name=rhl1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Taylor H.M., Grainger M., Holder A.A.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ430088; CAD33028.1; -.
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 142 AA; 17128 MW; 3DF9915D1941CD25 CRC64;

Query Match
Best Local Similarity 100.0%; Score 84; DB 2; Length 142;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 12
Db 120 HNNHNNHNNHNN 131

RESULT 4
O6Z1Z3
ID O6Z1Z3 PRELIMINARY; PRT; 362 AA.
AC O6Z1Z3.
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Transcription factor viviparous 1-like.
GN Name=B147B12.12;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriarctidaceae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005406; BABD3551.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003340; TF_B3.
DR PROSITE; PS50863; B3; 1.
SQ SEQUENCE 362 AA; 39941 MW; 70F577BAEEB4763D CRC64;

Query Match
Best Local Similarity 100.0%; Score 84; DB 2; Length 362;

```

Best Local Similarity 100.0%; Pred. No. 0.00087;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
| | | | | | | | | | | | | |
Db 49 HNNHNNHNNH 60

RESULT 5

Q7VAF5 PRELIMINARY; PRT; 362 AA.

AC 07VAF5; 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
GN Putative GTPase, G3E family.
OS OrderedLOCusNames=Pro1508;
OC Prochlorococcus marinus.
OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1219;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SBAG / CCMP 1375 / SS120;
RX MEDLINE=22810154; PubMed=12917486; DOI=10.1073/pnas.1733211100;
RA Dufresne A., Salanoubat M., Patenaky F., Artiguenave F., Axmann I.M.,
RA Barre V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,
RA Makarova K.S., Ostrowski M., Ozias S., Robert C., Rogozin I.B.,
RA Scallan D.J., Tandeau de Marsac N., Weissenbach J., Wincker P.,
RA Wolf Y.I., Hees W.R.;
RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
RT a nearly minimal oxyphototrophic genome."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025 (2003).
DR EMBL; AE017165; AAC00552.1; -
DR InterPro; IPR003495; COBw.
DR Pfam; PF02492; COBw; 1.
KW Complete proteome.
SQ SEQUENCE 362 AA; 41296 MW; 59CD085C7A58834D CRC64;

Query Match 100.0%; Score 84; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.00087;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
| | | | | | | | | | | | | |
Db 211 HNNHNNHNNH 222

RESULT 6

Q9HEX9 PRELIMINARY; PRT; 439 AA.

AC 09HEX9; 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
GN Putative a-factor pheromone receptor Ste3a.
GN Name=ste3;
OS Pneumocystis carinii.
OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OX NCBI_TaxID=4754;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21135666; PubMed=11238389;
RA Smilian A.G., Sesterhenn T., Tanaka R., Cushion M.T.;
RT "The ste3 pheromone receptor gene of Pneumocystis carinii is
RT surrounded by a cluster of signal transduction genes."
RL Genetics 157:991-1002 (2001).
DR EMBL; AF309805; AAG38536.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004933; F:mating-type a-factor pheromone receptor act. .; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR001546; Phtmea_receptor.

DR InterPro; IPR001499; STE3_GPCR.
DR Pfam; PF02076; STE3_1.
DR PRINTS; PR00899; GPCRSTE3.
DR PRINTS; PR00900; PHEROMONEAR.
KW Receptor.
SQ SEQUENCE 439 AA; 51337 MW; D545964E980A91B7 CRC64;

Query Match 100.0%; Score 84; DB 2; Length 439;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
| | | | | | | | | | | | | |
Db 385 HNNHNNHNNH 396

RESULT 7

Q6CIY7 PRELIMINARY; PRT; 441 AA.

AC 06CIY7; 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DE 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
GN Similarities with DERA0F16368g Debaryomyces hansenii.
GN ORFNames=YAL10F12287g;
OS Yarrowia lipolytica CLIB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genolevures;

RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marc C., Neuvéglise C., Tilia E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Badoir A., Barre V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confantolero F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Keszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicoud J.M., Nikolski M., Ozias S., Olier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Strub M.L., Suleau A.,
RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic I., Bojotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Galliard C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts."
RL Nature 430:35-44 (2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope; (JUL-2004) to the EMBL/genbank/DBJ databases.
RL Submitted (JUL-2004) to the EMBL/genbank/DBJ databases.
DR EMBL; CR382132; CAG78132.1; -
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
SQ SEQUENCE 441 AA; 49306 MW; BA8FDB190768959C CRC64;

Query Match 100.0%; Score 84; DB 2; Length 441;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
| | | | | | | | | | | | | |
Db 268 HNNHNNHNNH 279

RESULT 8

Q86IF5 PRELIMINARY; PRT; 1125 AA.

ID 086IF5
AC 086IF5;

```

DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidida; Dictyostelium.
NCBI_TaxID=4689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092922; PubMed=12092910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szatranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox B., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.,
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.",
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC117176; AAC52150.1; -.
DR HSRP; Q98935; I77C.
DR InterPro; IPR000198; RhOGAP.
DR InterPro; IPR008936; Rho GAP.
DR Pfam; PF00620; RhOGAP; 1.
DR SMART; SM00324; RhOGAP; 1.
DR PROSITE; PS50238; RHOGAP; 1.
KW Hypothetical protein.
SQ SEQUENCE 1125 AA; 124562 MW; 5F5E8CE48668113 CRC64;

Query Match 100.0%; Score 84; DB 2; Length 1125;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 12
DB 1091 HNNHNNHNNHNN 1102

RESULT 9
O7RSW8 PRELIMINARY; PRT; 2072 AA.
AC O7RSW8;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein (Fragment).
GN Name=PY00234;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlson J.M., Anguilo S.V., Suh B.B., Koij T.W., Pertea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shalom S.O., van Aken S.E., Riedmuller S.B., Feldlyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabli A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salberg S.L., Venter J.C., Fraser C.W., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.,
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii.",
RL Nature 419:512-519(2002).
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01000065; EAA21840.1; -.

```

```

KW Hypothetical protein.
FT NON TER 2072 2072
SQ SEQUENCE 2072 AA; 233638 MW; 8DA3BD8677336A31 CRC64;

Query Match 100.0%; Score 84; DB 2; Length 2072;
Best Local Similarity 100.0%; Pred. No. 0.005;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 12
DB 735 HNNHNNHNNHNN 746

RESULT 10
O7KF73 PRELIMINARY; PRT; 2969 AA.
AC O7KF73;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Normocyte-binding protein 1.
GN Name=NBPI;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB3;
RX MEDLINE=21590493; PubMed=11733572;
RA Rayner J.C., Vargas-Serrato E., Huber C.S., Galinski M.R.,
RA Barnwell J.W.;
RT "A Plasmodium falciparum homologue of Plasmodium vivax reticulocyte
RT binding protein (PvRBP) defines a trypan-resistant erythrocyte
RT invasion pathway.",
RL J. Exp. Med. 194:1571-1581(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HB3;
RA Rayner J.C., Barnwell J.W., Galinski M.R., Tran T., Huber C.S.,
RA Corredor V.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF411930; AAL38219.2; -.
DR InterPro; IPR011591; Botulinum.
DR ProDom; PD001963; Botulinum; 2.
SQ SEQUENCE 2969 AA; 357476 MW; EEOA4BA8F41B3F8 CRC64;

Query Match 100.0%; Score 84; DB 2; Length 2969;
Best Local Similarity 100.0%; Pred. No. 0.0072;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 12
DB 2831 HNNHNNHNNHNN 2842

RESULT 11
O8WRS4 PRELIMINARY; PRT; 2976 AA.
AC O8WRS4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Normocyte-binding protein 1.
GN Name=NBPI;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7G8;
RX MEDLINE=21590493; PubMed=11733572;
RA Rayner J.C., Vargas-Serrato E., Huber C.S., Galinski M.R.,
RA Barnwell J.W.;

```

RT "A Plasmodium falciparum homologue of Plasmodium vivax reticulocyte
RT binding protein (PvRBP1) defines a trypsin-resistant erythrocyte
RT invasion pathway."
RL J. Exp. Med. 194:1571-1581(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=7G8;
RA Rayner J.C., Barnwell J.W., Galinski M.R., Tran T., Huber C.S.,
RA Corredor V.;
RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF411933; AAL38222.2; -;
DR PRODOM; PD001963; Boculinum; 2.
SQ SEQUENCE 2976 AA; 358102 MW; C00F4D4A1556867C CRC64;

Query Match 100.0%; Score 84; DB 2; Length 2976;
Best Local Similarity 100.0%; Pred. No. 0.0072;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
ID |||||
AC 08WP99 PRELIMINARY; PRT; 2977 AA.
DB 2842 HNNHNNHNNHNN 2853

RESULT 12
Q8WP99 PRELIMINARY; PRT; 2977 AA.
ID 08WP99;
AC 08WP99;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE Normocyte-binding protein 1.
GN Name=NBPI;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_Taxid=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVO;
RX MEDLINE=21590493; PubMed=1173572;
RA Rayner J.C., Vargas-Serrato E., Huber C.S., Galinski M.R.,
RA Barnwell J.W.;
RT "A Plasmodium falciparum homologue of Plasmodium vivax reticulocyte
RT binding protein (PvRBP1) defines a trypsin-resistant erythrocyte
RT invasion pathway."
RL J. Exp. Med. 194:1571-1581(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVO;
RA Rayner J.C., Barnwell J.W., Galinski M.R., Tran T., Huber C.S.,
RA Corredor V.;
RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF411933; AAL38218.2; -;
DR PRODOM; PD001963; Boculinum; 1.
SQ SEQUENCE 2977 AA; 358535 MW; 0D548C9267DB34A1 CRC64;

Query Match 100.0%; Score 84; DB 2; Length 2977;
Best Local Similarity 100.0%; Pred. No. 0.0072;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
ID |||||
AC 07S0U3 PRELIMINARY; PRT; 417 AA.
DB 2839 HNNHNNHNNHNN 2850

RESULT 13
Q7S0U3 PRELIMINARY; PRT; 417 AA.
ID 07S0U3;
AC 07S0U3;
DT 01-MAR-2004 (TEMBLrel. 26, Created)
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU04619.1;
RP

OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_Taxid=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Puccell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Mearns M.,
RA Selltreimkoef C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothie G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysellis M., Mauceli E., Bielke C., Rudd S., Frisman D.,
RA Krysotova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cognoli C., Macino G., Catchside D., Li W., Pratt R.J., Omani S.A.,
RA Desouza C.C., Glaes L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebdole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AABX01000479; EAA28939.1; -;
DR GO; GO:0005634; C:nucleus; IEA;
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf_C2H2; 1.
DR PROSITE; PSS0157; ZINC FINGER C2H2 2; 1.
SQ SEQUENCE 417 AA; 45689 MW; AD225F345506F546 CRC64;

Query Match 92.9%; Score 78; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 11
ID |||||
AC MPT5_YEAST STANDARD; PRT; 834 AA.
DB 396 HNNHNNHNNHNN 406

RESULT 14
MPT5_YEAST STANDARD; PRT; 834 AA.
ID MPT5_YEAST
AC P39016;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Suppressor protein MPT5 (HTR1 protein).
GN Name=MPT5; Synonyms=HTR1, PUF5; OrderedLocNames=YGL178W;
GN ORFNames=BIC834;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Sakai A.;
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 70391;
RX MEDLINE=95147841; PubMed=7845352;
RA Kikuchi Y., Oka Y., Kobayashi M., Uesono Y., Toh-E A., Kikuchi A.;
RT "A new yeast gene, HTR1, required for growth at high temperature, is
RT needed for recovery from mating pheromone-induced G1 arrest."
RL Mol. Gen. Genet. 245:107-116(1994).
RN [3]
RP SEQUENCE FROM N.A.

```

RC STRAIN=5288c / FY1679;
RA MEDLINE=95397594; PubMed=766046;
RA Cogliavina M., Bertani I., Klma R., Zaccaria P., Bruschi C.V.;
RT "The DNA sequence of a 7941 bp fragment of the left arm of chromosome
RT VII of Saccharomyces cerevisiae contains four open reading frames
RT including the multicopy suppressor gene of the pop2 mutation and a
RT putative serine/threonine protein kinase gene."
RL Yeast 11:767-774(1995).
CC -1- FUNCTION: Multicopy suppressor of Pop2 mutation. Required for high
CC temperature growth.
CC -1- SIMILARITY: Contains 1 pum1lo-HD domain.
CC -1- SIMILARITY: Contains 1 pum1lo repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D26184; BAA05172.1; -
DR EMBL; D25541; BAA05024.1; -
DR EMBL; X83690; CAA58660.1; ALT_INIT.
DR EMBL; X83690; CAA58663.1; -
DR EMBL; Z72700; CAA96889.1; ALT_INIT.
DR PIR; S64195; S64195.
DR HSP; Q14671; 11B2.
DR Germonline; 141226; -.
DR SGD; S000003146; MPTS.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0003729; F:mRNA binding; IDA.
DR GO; GO:0007047; P:cell wall organization and biogenesis; IMP.
DR GO; GO:0001308; P:loss of chromatin silencing during replicat. .; IEP.
DR GO; GO:0000288; P:mRNA catabolism, deadenylation-dependent; IGI.
DR GO; GO:0006055; P:protein targeting; IMP.
DR GO; GO:0000321; P:re-entry into mitotic cell cycle after pher. .; IMP.
DR Interpro; IPR008939; ARM.
DR Interpro; IPR001313; Pumilio/Puf.
DR Pfam; PF00806; PUF; 8.
DR SMART; SM00025; Pumilio; 8.
DR PROSITE; PS50302; PUM; 8.
DR PROSITE; PS50303; PUM_HD; 1.
DR Repeat; RNA-binding.
KW DOMAIN
FT REPEAT 163 571 PUM-HD.
FT REPEAT 184 222 Pumilio 1.
FT REPEAT 223 258 Pumilio 2.
FT REPEAT 259 295 Pumilio 3.
FT REPEAT 300 337 Pumilio 4.
FT REPEAT 338 375 Pumilio 5.
FT REPEAT 376 413 Pumilio 6.
FT REPEAT 414 448 Pumilio 7.
FT REPEAT 478 514 Pumilio 8.
FT DOMAIN 622 628 Poly-Asn.
FT CONFLICT 777 834 RMOYQTEGAMDSLSMKSOHIGQSPYQNMVMSNASTISM
FT PAMTARTSDDELQFTLP -> QMNTPTPRIML (in Ref.
SQ SEQUENCE 834 AA; 92828 MW; 8391EB04764D47EB CRC64;
Query Match 92.9%; Score 78; DB 1; Length 834;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HNNHNNHNNH 11
DB 609 HNNHNNHNNH 619

```

```

DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Transposable element activator hypothetical 12 kDa protein (AC 12 kDa
DE protein).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Kunze R., Stochaj U., Laufs J., Starlinger P.;
RT "transcription of transposable element Activator (Ac) of Zea mays L.";
RL EMBL J. 6:1555-1563(1987).
CC -1- MISCELLANEOUS: This protein is coded by the transposable maize
CC controlling element "Activator" (Ac), which is able to activate
CC chromosome breakage at a specific location.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X05424; CAA29006.1; -
DR PIR; T02917; T02917.
DR MaizEDB; 69192; -.
KW Hypothetical protein; Transposable element.
SQ SEQUENCE 102 AA; 12406 MW; 0A7B6ED251F9E1 CRC64;
Query Match 90.5%; Score 76; DB 1; Length 102;
Best Local Similarity 91.7%; Pred. No. 0.0027;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 HNNHNNHNNH 12
DB 20 HNNHNNHNNH 31

```

```

RESULT 16
ID Q81516 PRELIMINARY; PRT; 501 AA.
AC Q81516;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein.
GN ORFNames=PF11100;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Hall N., Nelson K.B., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,
RA Pereira M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carninci D.J., Holtman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrett B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Hyman R.W., Fung E., Conway A., Kurdi O., Mac J., Miranda M.,
RA Nakao B., Rowley D., Tamaki T., Wang F., Davis R.W.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

```


DR EMBL; AE014847; AAN36306.1; -.
 KW Hypochemical protein.
 SQ SEQUENCE 501 AA; 60281 MW; 54663449C7C67AC3 CRC64;

Query Match 90.5%; Score 76; DB 2; Length 501;
 Best Local Similarity 100.0%; Pred. No. 0.013;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NNNNNNNNNN 12
 |||||
 Db 228 NNNNNNNNNN 238

RESULT 17

NEUR DROVI STANDARD; PRT; 747 AA.
 ID NEUR DROVI
 AC 024726;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Neutralized protein.
 GN Name=neur; Synonym=neu;
 OS Drosophila virilis (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxId=7244;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95095077; PubMed=8001814;
 RA Zhou L., Boulianne G.L.;
 RT "Comparison of the neutralized genes of Drosophila virilis and D.
 melanogaster.";
 RL Genome 37:840-847(1994).
 CC -1- FUNCTION: Involved in neurogenesis. Interacts with other
 neurogenic proteins in the specification of the neuroblast versus
 epidermal cell fate.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- SIMILARITY: Contains 2 NEUZ domains.
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U12593; AAB60619.1; -.
 CC EMBL; U12591; AAB60619.1; JOINED.
 CC EMBL; U12592; AAB60619.1; JOINED.
 DR FLYBASE; FBgn0013312; Dv1r\neur.
 DR GO; GO:0005886; C:plasma membrane; ISS.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; ISS.
 DR GO; GO:0007456; P:eye morphogenesis (sensu Drosophila); ISS.
 DR GO; GO:0007498; P:mesoderm development; ISS.
 DR GO; GO:0007423; P:sensory organ development; ISS.
 DR GO; GO:0016360; P:sensory organ precursor cell fate determina. . .; ISS.
 DR InterPro; IPR006573; Neu_2.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF00977; Zf-C3HC4; 2.
 DR SMART; SM00588; NEUZ; 2.
 DR SMART; SM00184; RING; 1.
 DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
 DR PROSITE; PSS0089; ZF_RING_2; 1.
 KW DNA-binding; Neurogenesis; Nuclear protein; Repeat; Zinc-finger.
 FT DOMAIN 95 216 NEUZ 1.
 FT ZN FING 357 480 NEUZ 2.
 FT DOMAIN 694 735 RING-type.
 FT ZN FING 24 35 His-rich.
 FT DOMAIN 257 300 His-rich.

FT DOMAIN 530 544 Ala-rich.
 SQ SEQUENCE 747 AA; 81982 MW; C8E72569D6FDCA11 CRC64;

Query Match 88.1%; Score 74; DB 1; Length 747;
 Best Local Similarity 83.3%; Pred. No. 0.036;

Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNNNNNNNN 12
 |||||
 Db 25 HNNNNNNNNN 36

RESULT 18

Q8MCQO PRELIMINARY; PRT; 1544 AA.
 ID Q8MCQO
 AC Q8MCQO;
 DT 01-OCT-2002 (T-EMBLrel. 22, Created)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE S013096p.
 GN Name=Pdel1; Synonym=CG10231;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacle J., Paragas V., Park S.,
 RA Patel S., Phouenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY122262; AAM52774.1; -.
 DR FLYBASE; FBgn0032686; Pdel1.
 DR GO; GO:0004114; F:3',5'-cyclic-nucleotide phosphodiesterase a. . .; IEA.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR003018; GAF.
 DR InterPro; IPR003607; Met_phos_hydro.
 DR InterPro; IPR002073; PDBase.
 DR Pfam; PF01590; GAF; 2.
 DR Pfam; PF00233; PDBase_1; 1.
 DR PRINTS; PRO0387; PDISTERASE1.
 DR SMART; SM00065; GAF; 2.
 DR SMART; SM00471; Hdc; 1.
 DR PROSITE; PS00126; PDASE_1; 1.
 DR SEQUENCE 1544 AA; 171689 MW; F5254E0CCB3D065 CRC64;

Query Match 88.1%; Score 74; DB 2; Length 1544;
 Best Local Similarity 83.3%; Pred. No. 0.074;

Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNNNNNNNN 12
 |||||
 Db 1423 HNNNNNNNNN 1434

RESULT 19

O43988 PRELIMINARY; PRT; 800 AA.
 ID O43988
 AC O43988;
 DT 01-JUN-1998 (T-EMBLrel. 06, Created)
 DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE Homeobox-containing protein Marial (Fragment).
 GN Name=marial;
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
 NCBI_TaxId=44689;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=KAX3;
 RA Han Z., Firtel R.A.;
 RT "The homeobox-containing gene *Warrior* regulates anterior-posterior
 RT patterning and cell-type homeostasis in *Dictyostelium*.";
 RL Development 0:0-0(1998).
 CC -1- SIMILARITY: Contains 1 histidine kinase domain.
 DR EMBL: AF036170; AAB92245.1; -.
 DR HSSP: P09959; ISW6.
 DR DictyBase: DDB0191441; wara.
 DR GO: GO:0005634; C:nucleus; IEA.
 DR GO: GO:0003700; F:transcription factor activity; IEA.
 DR GO: GO:0006555; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR001356; Homeobox.
 DR InterPro: IPR009057; Homeodomain_like.
 DR Pfam: PF00023; ANK; 9.
 DR Pfam: PF00046; Homeobox; 1.
 DR PRINTS: PRO1415; ANKYRIN.
 DR PRINTS: PRO0024; HOMEBOX.
 DR ProDom: PD000010; Homeobox; 1.
 DR SMART: SM00248; ANK; 8.
 DR SMART: SM00389; HOX; 1.
 DR PROSITE: PS50088; ANK_REPEAT; 8.
 DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE: PS00027; HOMEBOX_1; 1.
 DR PROSITE: PS50071; HOMEBOX_2; 1.
 KW ANK repeat; DNA-binding; Homeobox; Nuclear protein.
 FT NON TER 800
 SQ SEQUENCE 800 AA; 88723 MW; 7FACC2C0984DB4F CRC64;

Query Match 87.5%; Score 73.5; DB 2; Length 800;
 Best Local Similarity 92.3%; Pred. No. 0.045; Mismatches 0; Indels 1; Gaps 1;
 Matches 12; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Oy 1 HNNHNNHNNHNN 12
 Db 303 HNNHNNHNNHNN 315

RESULT 20
 Q23901
 ID Q23901 PRELIMINARY; PRT; 1670 AA.
 AC Q23901;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Histidine kinase.
 GN Name=doka;
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 OX NCBI_TaxId=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX2;
 RA MEDLINE=96324396; PubMed=8670893;
 RA Schuster S.C., Noegel A.A., Oehme F., Gerisch G., Simon M.I.;
 RT "The hybrid histidine kinase Doka is part of the osmotic response
 RT system of *Dictyostelium*.";
 RL EMBL J. 15:3880-3889(1996).
 CC -1- SIMILARITY: Contains 1 histidine kinase domain.
 DR EMBL: X96869; CAA65612.1; -.
 DR PIR: S71628; S71628.
 DR HSSP: O9A514; IMVY.
 DR DictyBase: DDB0185194; doka.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0003677; F:DNA binding; IEA.
 DR GO: GO:0016301; F:kinase activity; IEA.
 DR GO: GO:0000156; F:two-component response regulator activity; IEA.
 DR GO: GO:0000155; F:two-component sensor molecule activity; IEA.
 DR GO: GO:0007600; P:sensory perception; IEA.
 DR GO: GO:0000160; P:two-component signal transduction system (p. . .; IEA.

DR InterPro: IPR003594; ATPbind_ATPase.
 DR InterPro: IPR004358; Bact_sens_pr_C.
 DR InterPro: IPR001128; Cytochrome_P450.
 DR InterPro: IPR005467; His_kinase.
 DR InterPro: IPR003661; His_kinase_N.
 DR InterPro: IPR000014; PAS.
 DR InterPro: IPR001789; Response_reg.
 DR Pfam: PF02518; HATPase_c; 1.
 DR Pfam: PF00512; HisKA_I.
 DR Pfam: PF00072; Response_reg; 1.
 DR PRINTS: PRO0344; BCTRSENSOR.
 DR SMART: SM00387; HATPase_c; 1.
 DR SMART: SM00388; HisKA; 1.
 DR SMART: SM00091; PAS; 2.
 DR SMART: SM00448; REC; 1.
 DR TIGR/Pfam: TIGR00229; sensory_box; 1.
 DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
 DR PROSITE: PS50109; HIS_KIN; 1.
 DR PROSITE: PS50110; RESPONSE_REGULATORY; 1.
 KW kinase; Phosphorylation; Sensory transduction.
 SQ SEQUENCE 1670 AA; 186391 MW; 85C76DEE847276B5 CRC64;

Query Match 87.5%; Score 73.5; DB 2; Length 1670;
 Best Local Similarity 92.3%; Pred. No. 0.093; Mismatches 0; Indels 1; Gaps 1;
 Matches 12; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Oy 1 HNNHNNHNNHNN 12
 Db 576 HNNHNNHNNHNN 588

RESULT 21
 Q869S5
 ID Q869S5 PRELIMINARY; PRT; 1671 AA.
 AC Q869S5;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Similar to Dictyostelium discoideum (Slime mold). Histidine
 DE kinase.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 OX NCBI_TaxId=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
 RA Gloeckner G., Eichinger U., Szallasi Z., Pachabati J., Dear P.,
 RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
 RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
 RT "Sequence and analysis of chromosome 2 of *Dictyostelium discoideum*.";
 RL Nature 418:79-85(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Baumgart C.;
 RT Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 histidine kinase domain.
 DR EMBL: AC116957; AA052477.1; -.
 DR HSSP: O9A514; IMVY.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0003677; F:DNA binding; IEA.
 DR GO: GO:0016301; F:kinase activity; IEA.
 DR GO: GO:0000155; F:two-component response regulator activity; IEA.
 DR GO: GO:0000156; F:two-component sensor molecule activity; IEA.
 DR GO: GO:0007600; P:sensory perception; IEA.
 DR InterPro: IPR003594; ATPbind_ATPase.
 DR InterPro: IPR004358; Bact_sens_pr_C.
 DR InterPro: IPR001128; Cytochrome_P450.
 DR InterPro: IPR005467; His_kinase.
 DR InterPro: IPR003661; His_kinase_N.

```
DR InterPro: IPR00014; PAS.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF02518; HATPase_c_1.
DR Pfam: PF00512; Hleka; 1.
DR PRINTS: PR00712; Response_reg; 1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR ProDom: PD000039; Response_reg; 1.
DR SMART: SM00387; HATPase_c_1.
DR SMART: SM00388; Hleka; 1.
DR SMART: SM00091; PAS; 2.
DR SMART: SM00448; REC; 1.
DR TIGRFAKE: TIGR00229; sensory_box; 1.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
DR PROSITE: PS50109; HIS_KIN; 1.
DR PROSITE: PS50110; RESPONSE_REGULATORY; 1.
DR KINase: Phosphorylation; Sensory transduction.
SQ SEQUENCE 1671 AA; 186135 MW; ED349E0613B43345 CRC64;

Query Match      87.5%; Score 73.5; DB 2; Length 1671;
Best Local Similarity 92.3%; Pred. No. 0.093;
Matches 12; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 HNNHNNHNNHNN 12
   |||||
DB 577 HNNHNNHNNHNN 589

RESULT 22
ID Q815M3 PRELIMINARY; PRT; 1061 AA.
AC Q815M3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PF01895C;
OS Plasmodium falciparum (Isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallow S.J., Sub B., Peterson J., Anguoli S.,
RA Perlea M., Allen J., Selengut J., Haft D., Mather M.W., Valdivia A.B.,
RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrall B.,
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RT Nature 419:498-511(2002).
RT [2]
RP SEQUENCE FROM N.A.
RX Hyman R.W., Fung E., Conway A., Kurdi O., Mao J., Miranda M.,
RA Nakao B., Rowley D., Tamaki T., Wang F., Davis R.W.,
RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AEO14846; AAN36267.1; -
DR GO; GO:0005525; P:GTP binding; IEA.
DR GO; GO:0003924; P:GTPase activity; IEA.
DR InterPro: IPR001401; DYNAMIN.
DR PRINTS; PR00195; DYNAMIN.
KW Hypothetical protein.
SQ SEQUENCE 1061 AA; 122755 MW; F401565EFBAC4C7 CRC64;

Query Match      84.5%; Score 71; DB 2; Length 1061;
Best Local Similarity 83.3%; Pred. No. 0.13;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
   |||||
DB 577 HNNHNNHNNHNN 589

DB 657 HNNHNNHNNHNN 668

RESULT 23
ID Q7ROA5 PRELIMINARY; PRT; 1210 AA.
AC Q7ROA5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Vacuolar protein sorting homolog r-vps33a.
GN Name=PY01196;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=33239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlson J.M., Anguoli S.V., Sub B.B., Kool J.T.W., Perlea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallow S.J., van Aken S.E., Riedmiller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabli A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Valdivia A.B.,
RA van Lin L.H., Jansz J.C., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RT Nature 419:512-519(2002).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01000315; BAA20486.1; -
DR GO; GO:0008565; P:protein transporter activity; IEA.
DR GO; GO:0016192; P:vesicle-mediated transport; IEA.
DR InterPro: IPR001619; Sec1-like.
DR Pfam: PF00995; Sec1; 1.
SQ SEQUENCE 1210 AA; 142560 MW; 5371B06BC370FCCE CRC64;

Query Match      84.5%; Score 71; DB 2; Length 1210;
Best Local Similarity 90.9%; Pred. No. 0.14;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HNNHNNHNNHNN 12
   |||||
DB 334 HNNHNNHNNHNN 344

RESULT 24
ID Q9P826 PRELIMINARY; PRT; 233 AA.
AC Q9P826;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein (Fragment).
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21155066; PubMed=11231556;
RA De Backer M.D., Nelissen B., Logghe M., Viaene J., Loomen I.,
RA Vandewinck S., de Hoogt R., Dewaele S., Simons F.A., Verhasselt P.,
RA Vanhoof G., Contreras R., Luyten W.H.;
RT "An antisense-based functional genomics approach for identification of
RT genes critical for growth of Candida albicans."
RT Nat. Biotechnol. 19:212-213(2001).
DR EMBL; AJ390515; CAB77653.1; -
```

DR GO; GO:0000175; F:3'-5'-exoribonuclease activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0006396; P:RNA processing; IEA.
 DR InterPro; IPR001247; 3 EXORBASE.
 DR InterPro; IPR010400; DUF1000.
 DR Pfam; PF06201; DUF1000; 1.
 DR Hypothetical protein.
 FT NON_TER
 SQ SEQUENCE 233 AA; 26266 MW; 74707321AF895CE7 CRC64;

Query Match 83.3%; Score 70; DB 2; Length 233;
 Best Local Similarity 83.3%; Pred. No. 0.037;
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
 |||||
 DB 8 HNNHNNHNNHNN 19

RESULT 25

CAR3_DICDI STANDARD; PRT; 490 AA.
 AC P35352;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Cyclic AMP receptor 3.
 GN Name-carc; Synonyms=car3;
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX3;
 RX MEDLINE=93170666; PubMed=8382181;
 RA Johnson R.L., Saxe C.L. III, Gollig R., Kimmel A.R., Devreotes P.N.;
 RT "Identification and targeted gene disruption of car3, a cAMP receptor
 subunit expressed during multicellular stages of Dictyostelium
 development.";
 RL Genes Dev. 7:273-282(1993).

CC -1- FUNCTION: Receptor for cAMP. Coordinates the aggregation of
 individual cells into a multicellular organism and regulates the
 expression of a large number of developmentally regulated genes.
 CC The activity of this receptor is mediated by G proteins.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- DEVELOPMENTAL STAGE: Induced at early aggregation (6 hrs) and
 maximally expressed at the mound stage (9-12 hrs), level of
 expression peaks again during the slug stage (18 hrs) and declines
 at culmination.
 CC -1- PTM: Carboxyl-terminal Ser or Thr residues may be phosphorylated.
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 5 family.
 CC -1- CAUTION: It is uncertain whether Met-1 or Met-15 is the initiator.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 or send an email to license@ebi.ac.uk).

CC EMBL; S55235; AAB55437.1; -
 DR DictyBase; DDB0003726; CARC.
 DR InterPro; IPR000848; GPCR_CAMP.
 DR InterPro; IPR000832; GPCR_secretin.
 DR Pfam; PF05462; Dicty_CAR; 1.
 DR PRINTS; PR00247; GPCRAMP.
 DR PROSITE; PS00261; G_PROTEIN_RECEP_F2_4; 1.
 DR G-protein coupled receptor; Glycoprotein; Multigene family;
 KW Phosphorylation; Transmembrane.
 FT DOMAIN 1 23 Extracellular (Potential).
 FT TRANSMEM 43 1 (Potential).
 FT DOMAIN 44 57 Cytoplasmic (Potential).

FT TRANSMEM 58 78 2 (Potential).
 FT DOMAIN 79 94 Extracellular (Potential).
 FT TRANSMEM 95 120 3 (Potential).
 FT DOMAIN 121 131 Cytoplasmic (Potential).
 FT TRANSMEM 132 150 4 (Potential).
 FT DOMAIN 151 173 Extracellular (Potential).
 FT TRANSMEM 174 192 5 (Potential).
 FT DOMAIN 193 216 Cytoplasmic (Potential).
 FT TRANSMEM 217 235 6 (Potential).
 FT DOMAIN 236 246 Extracellular (Potential).
 FT TRANSMEM 247 271 7 (Potential).
 FT DOMAIN 272 490 Cytoplasmic (Potential).
 FT MOD_RES 204 204 Phosphoserine (by PKA) (Potential).
 FT DOMAIN 303 439 Aen-rich.
 FT DOMAIN 399 427 Poly-Asn.
 SQ SEQUENCE 490 AA; 56161 MW; A28BA33408626153 CRC64;

Query Match 83.3%; Score 70; DB 1; Length 490;
 Best Local Similarity 83.3%; Pred. No. 0.078;
 Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
 |||||
 DB 325 HNNHNNHNNHNN 336

RESULT 26

Q869W0 PRELIMINARY; PRT; 942 AA.
 AC Q869W0;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Similar to Dictyostelium discoideum (Slime mold). homeobox-containing
 protein.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
 RA Gloeckner G., Eichinger L., Szatranski K., Pechabat J., Dear P.,
 RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
 RA Tunngal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
 RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
 RL Nature 416:79-85(2002).

CC -1- FUNCTION: Receptor for cAMP. Coordinates the aggregation of
 individual cells into a multicellular organism and regulates the
 expression of a large number of developmentally regulated genes.
 CC The activity of this receptor is mediated by G proteins.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- DEVELOPMENTAL STAGE: Induced at early aggregation (6 hrs) and
 maximally expressed at the mound stage (9-12 hrs), level of
 expression peaks again during the slug stage (18 hrs) and declines
 at culmination.
 CC -1- PTM: Carboxyl-terminal Ser or Thr residues may be phosphorylated.
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 5 family.
 CC -1- CAUTION: It is uncertain whether Met-1 or Met-15 is the initiator.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 or send an email to license@ebi.ac.uk).

CC EMBL; AC116305; AAO52354.1;
 DR EMBL; AC116305; AAO52354.1;
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR009057; Homeobox_like.
 DR Pfam; PF00046; Homeobox; 1.
 DR Prodom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOK; 1.
 DR PROSITE; PS00027; HOMEBOX_1; UNKNOWN_1.
 DR PROSITE; PS00711; HOMEBOX_2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 SQ SEQUENCE 942 AA; 107111 MW; 33A69BA8616A493 CRC64;

Query Match 83.3%; Score 70; DB 2; Length 942;
 Best Local Similarity 83.3%; Pred. No. 0.15;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
 |||||

Db 383 HNNHNNHNNH 394

RESULT 27

08T1A1 PRELIMINARY; PRT; 72 AA.

AC 08T1A1; 01-JUN-2002 (T-EMBLrel. 21, Created)

DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)

DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

DE Hypothetical protein

OS Dictyostelium discoideum (Slime mold).

OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

OX NCBI_TaxId=44689;

11

RP SEQUENCE FROM N.A.

RC STRAIN=AX4;

RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;

RA Gloeckner G., Eichinger L., Szatranski K., Pachebat J., Dear P., Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K., Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;

RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";

RL Nature 418:79-85(2002).

12

RP SEQUENCE FROM N.A.

RC STRAIN=AX4;

RA Baumgart C.;

RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL, Acc16956; AAM08755.1; -.

KM Hypothetical protein.

SQ SEQUENCE 72 AA; 8867 MW; 27CDA105612PDCA CRC64;

Query Match 82.1%; Score 69; DB 2; Length 72;

Best Local Similarity 75.0%; Pired. No. 0.015;

Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12

Db 56 HNNHNNHNNH 67

RESULT 28

09VJ79 PRELIMINARY; PRT; 1365 AA.

AC 09VJ79; 01-MAY-2000 (T-EMBLrel. 13, Created)

DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)

DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)

DE CG10231-PA.

GN Name=Pde11; ORFNames=CG10231;

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxId=7227;

11

RP SEQUENCE FROM N.A.

RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang M., Pfeiffer B.D., Brandon R.C., Rogers Y.H., Blazek R.G., Chame M., Pfeiffer B.D., Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L., Abril J.F., Agbayani A., An H.J., Andrews-Pfankuch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotier P., Burts K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Laslo P., Lei Y., Levitsky A.A., Li Z., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Moberly C., Morris J., Mostreli A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

12

RP SEQUENCE FROM N.A.

RX MEDLINE=22426065; PubMed=12537568;

RA Celinker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A., Patel S., Adams M., Chame M., Dugan S.P., Frise E., Hodgson A., George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R., Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodegren E.J., Swirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C., Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;

RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila melanogaster euchromatic genome sequence.";

RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

13

RP SEQUENCE FROM N.A.

RX MEDLINE=22426070; PubMed=12537573;

RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Swirskas R., Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M., Ashburner M., Celinker S.E.;

RT "The transposable elements of the Drosophila melanogaster euchromatin: a genomic perspective.";

RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).

14

RP SEQUENCE FROM N.A.

RX MEDLINE=22426069; PubMed=12537572;

RA Misra S., Crosby M.A., Mungall C.U., Matthews B.B., Campbell K.S., Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E., Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P., Bettencourt B.R., Celinker S.E., de Grey A.D., Drysdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M., Lewis S.E.;

RT "Annotation of the Drosophila melanogaster euchromatic genome: a systematic review.";

RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).

15

RP SEQUENCE FROM N.A.

RX FlyBase; 01-JUN-2002 (T-EMBLrel. 21, Created)

DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)

DE Pde11

GN Name=Pde11; ORFNames=CG10231;

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxId=7227;

11

RP SEQUENCE FROM N.A.

RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang M., Pfeiffer B.D., Brandon R.C., Rogers Y.H., Blazek R.G., Chame M., Pfeiffer B.D., Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L., Abril J.F., Agbayani A., An H.J., Andrews-Pfankuch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotier P., Burts K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,

```

DR Pfam; PF01590; GAF; 2.
DR Pfam; PF00233; PDase_I; 1.
DR PRINTS; PR00387; PDSESTERASE1.
DR SMART; SM00065; GAF; 2.
DR SMART; SM00471; HDC; 1.
DR PROSITE; PS00126; PDASE_I; 1.
SQ SEQUENCE 1365 AA; 15173 MW; B0562EFDF5012E56 CRC64;

Query Match      82.1%; Score 69; DB 2; Length 1365;
Best Local Similarity 75.0%; Pred. No. 0.29;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHHNNHHNN 12
Db 1244 HHHHHSHSHNN 1255

RESULT 29
Q8ILV0 PRELIMINARY; PRT; 2770 AA.
AC Q8ILV0;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PF14_0143;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA GARDNER M.J., Hall N., Fung E., White C., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.B., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallow S.J., Sun B., Peterson J., Angiuoli S.,
RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Mather M.W., Vaidya A.B.,
RA Venter J.C., Canciani D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrett B.;
RT "Genome sequence of the human malaria parasite Plasmodium
falciparum."
RL Nature 419:498-511(2002).
DR EMBL; AB014818; AAN36755.1; -.
DR InterPro; IPR004147; ABC_1.
DR Pfam; PF03109; ABC1; 1.
KW Hypothetical protein.
SQ SEQUENCE 2770 AA; 327405 MW; 04F06D726B630AD6 CRC64;

Query Match      82.1%; Score 69; DB 2; Length 2770;
Best Local Similarity 75.0%; Pred. No. 0.59;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHHNNHHNN 12
Db 1219 HHHHHHHHHNNH 1230

RESULT 30
Q9VZX2 PRELIMINARY; PRT; 753 AA.
AC Q9VZX2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE CG9973-PA.
GN ORFNames=CG9973;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.P.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabot G.L.,
RA April J.P., Agbayani A., An H.J., Andrews-Pfankuch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Kocha S., Dunkov B.C., Dunn P.,
RA Durbin K.U., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.H., Ibeagwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Misha N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard R., Paclob J.M.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Spadling A.C., Stapleton M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Simpson M., Skupski M.P., Smith T.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodard J., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao X., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs J.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).

[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Ceiniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Friese E., Hodgeson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Paclob J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
RA Patel S., Friese E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Ceiniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomic perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).

[4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Hirste S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield B.U., Bayraktaroglu L., Berman B.P.,
RA Betencourt B.R., Ceiniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a

```

RT systematic review.";
 RN Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RG Flybase:
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RN SEQUENCE FROM N.A.
 RG Flybase:
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE003476; AAF47693.2; -.
 DR Flybase: Fgn0035378; CG9973.
 DR GO: GO:0003777; F:DNA binding; IEA.
 DR GO: GO:0008270; F:zinc ion binding; IEA.
 DR InterPro: IPR01051; RmlC like cupin.
 DR InterPro: IPR002857; Znf_CXXC.
 DR Pfam: PR02008; zf-CXXC; T.
 SQ SEQUENCE 753 AA; 7998 MW; F5E6CD1F030829A2 CRC64;

Query Match 81.5%; Score 68.5; DB 2; Length 753;
 Best Local Similarity 84.6%; Pred. No. 0.19;
 Matches 11; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 HNNHNNNNNNH 12
 DB 622 HNNHNNNNNNH 634

RESULT 31

Q878F2 PRELIMINARY; PRT; 136 AA.

DT 01-JUN-2002 (TREMblrel. 21, Created)

DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)

DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)

DE Rhl protein (Fragment).

GN Name=rhl;

OS Plasmodium falciiparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.

OX NCBI_TaxID=5893;

RN [1]

RP SEQUENCE FROM N.A.

RA Taylor H.M., Grainger M., Holder A.A.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ430087; CAD23027.1; -.

FT NON_TER

FT 136

FT 136

FT 136

FT 136

FT 136

FT 136

FT 136

FT 136

FT 136

FT 136

FT 136

FT 136

FT 136

RC STRAIN=ATCC VR-111;
 RX PubMed=14671117; DOI=10.1128/JVI.78.1.353-366.2004;
 RA Tulman E.R., Afonso C.L., Lu Z., Zeak L., Kutish G.F., Rock D.L.;
 RT "The genome of canarypox virus."
 RL J. Virol. 78:353-366(2004).
 DR EMBL: AF18871; AAR83507.1; -.
 DR InterPro: IPR002400; GP_CysKnot.
 DR InterPro: IPR001839; TGFb.
 DR Pfam: PF00019; TGF beta; 1.
 DR PRINTS: PR00438; GFCYSKNOT.
 DR ProDom: PD000357; TGFb; 1.
 DR SMART: SM00204; TGFb; 1.
 SQ SEQUENCE 358 AA; 41528 MW; 7A2FB2CEE1DFE51D CRC64;

QY 2 NNNNNNNNNH 12
 DB 237 NNNNNNNNNH 247

RESULT 33

Q869Z3 PRELIMINARY; PRT; 584 AA.

DT 01-JUN-2003 (TREMblrel. 24, Created)

DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)

DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)

DE Similar to Fusobacterium nucleatum (Subsp. nucleatum). GlutamyI-cRNA synthetase (EC 6.1.1.17).

OS Dictyostelium discoideum (Slime mold).

OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

OX NCBI_TaxID=44689;

RN [1]

RP SEQUENCE FROM N.A.

RA Baumgart C.;

RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC116925; AAO52252.1; -.

DR HSPF: P27000; 1U09.

DR GO: GO:0005737; C:cytoplasm; IEA.

DR GO: GO:0005524; F:ATP binding; IEA.

DR GO: GO:0004818; F:glutamate-cRNA ligase activity; IEA.

DR GO: GO:0016874; F:ligase activity; IEA.

DR GO: GO:0006424; P:glutamyI-cRNA aminoacylation; IEA.

DR GO: GO:0006412; P:protein biosynthesis; IEA.

DR InterPro: IPR004527; GltX_dact.

DR InterPro: IPR000924; Glu_cRNA-synt_1c.

DR InterPro: IPR008925; cRNA-synt bind.

DR InterPro: IPR001412; cRNA-synt_1.

DR Pfam: PR00749; tRNA-synt_1c; 1.

DR PRINTS: PR00987; TRNASYNTHGU.

DR TIGRPFam: TIGR00464; GltX_dact; 1.

DR PROSITE: PS00178; AA_TRNA_LIGASE_1; 1.

QY 81.0%; Score 68; DB 2; Length 584;
 Best Local Similarity 90.9%; Pred. No. 0.1;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNNNNNNNNH 12
 DB 237 NNNNNNNNNH 247

RESULT 33

Q869Z3 PRELIMINARY; PRT; 584 AA.

DT 01-JUN-2003 (TREMblrel. 24, Created)

DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)

DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)

DE Similar to Fusobacterium nucleatum (Subsp. nucleatum). GlutamyI-cRNA synthetase (EC 6.1.1.17).

OS Dictyostelium discoideum (Slime mold).

OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

OX NCBI_TaxID=44689;

RN [1]

RP SEQUENCE FROM N.A.

RA Baumgart C.;

RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC116925; AAO52252.1; -.

DR HSPF: P27000; 1U09.

DR GO: GO:0005737; C:cytoplasm; IEA.

DR GO: GO:0005524; F:ATP binding; IEA.

DR GO: GO:0004818; F:glutamate-cRNA ligase activity; IEA.

DR GO: GO:0016874; F:ligase activity; IEA.

DR GO: GO:0006424; P:glutamyI-cRNA aminoacylation; IEA.

DR GO: GO:0006412; P:protein biosynthesis; IEA.

DR InterPro: IPR004527; GltX_dact.

DR InterPro: IPR000924; Glu_cRNA-synt_1c.

DR InterPro: IPR008925; cRNA-synt bind.

DR InterPro: IPR001412; cRNA-synt_1.

DR Pfam: PR00749; tRNA-synt_1c; 1.

DR PRINTS: PR00987; TRNASYNTHGU.

DR TIGRPFam: TIGR00464; GltX_dact; 1.

DR PROSITE: PS00178; AA_TRNA_LIGASE_1; 1.

QY 1 HNNHNNHNNH 12
 DB 41 HNNHNNHNNH 52

RESULT 34

Q78953 PRELIMINARY; PRT; 604 AA.

AC Q78953; 01-MAR-2004 (TREMBlrel. 26, Created)
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
 DE Hypothetical protein.
 GN Name=NCU07262.1;
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 RN [1] NCBI_TaxID=5141;
 RP SEQUENCE FROM N.A.

RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
 RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
 RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
 RA Seltremikoff C.P., Kinsey J.A., Brun E.L., Zelter A., Schulte U.,
 RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
 RA Roy A., Foley K., Naylor J., Thoman N., Barrett R., Gnerre S.,
 RA Kamal M., Kamysella M., Mauceli E., Bielke C., Rudd S., Friesman D.,
 RA Kryofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
 RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
 RA Varden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
 RA Cogni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
 RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.,
 RA "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
 RL Nature 0.0-0(2003).
 CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL; AABX01000218; EAA32880.1; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008324; F:cation transporter activity; IEA.
 DR GO; GO:0006612; P:cation transport; IEA.
 DR InterPro; IPR002524; Cation_efflux.
 DR InterPro; IPR002395; Kinogen.
 DR Pfam; Pf01545; Cation_efflux; 1.
 DR PRINTS; PR00334; KININGEN.
 DR TIGRfams; TIGR01297; CDF; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 604 AA; 66355 MW; 1D9CAF3D3EE63CF CRC64;

Query Match 81.0%; Score 68; DB 2; Length 604;
 Best Local Similarity 75.0%; Pred. No. 0.17;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
 DB 97 HNNHNNHNNH 108

RESULT 35

Q8WR55 PRELIMINARY; PRT; 2965 AA.

AC Q8WR55; 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DE Normocyte-binding protein 1.
 GN Name=NBPL1;
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 RN [1] NCBI_TaxID=5833;

RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=Malayan Camp K-;
 RX MEDLINE=2130493; PubMed=11733572;
 RA Rayner J.C., Vargas-Serrato E., Huber C.S., Galinski M.R.,
 RA Barnwell J.W.;
 RT "A Plasmodium falciparum homolog of Plasmodium vivax reticulocyte
 RT binding protein (PyRBP1) defines a trypan-resistant erythrocyte
 RL invasion pathway".
 RL J. Exp. Med. 194:1571-1581(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Malayan Camp K-;
 RA Rayner J.C., Barnwell J.W., Galinski M.R., Tran T., Huber C.S.,
 RA Cordeiro V.;
 RL Submitted (Aug-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF411932; AAL38221.2;
 DR Prodom; PD001963; Botulinum; 2.
 SQ SEQUENCE 2965 AA; 356914 MW; 98B077462826A8FE CRC64;

Query Match 81.0%; Score 68; DB 2; Length 2965;
 Best Local Similarity 83.3%; Pred. No. 0.86;
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
 DB 2835 HNNHNNHNNH 2846

RESULT 36

Q6ALP7 PRELIMINARY; PRT; 89 AA.

AC Q6ALP7; 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DE Histidine-rich glycoprotein.
 OS Euploetes vanus.
 OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Hypotrichia;
 OC Euploetida; Euploetidae; Euploetes.
 RN [1] NCBI_TaxID=5939;
 RP SEQUENCE FROM N.A.
 RA Apel A.K., Hankeln T., Schmidt E.R.;
 RT "Survey sequencing of Euploetes vanus macronuclear genes.";
 RL Submitted (Jul-2003) to the EMBL/Genbank/DBJ databases.
 DR EMBL; A1698345; CAH04400.1; -;
 SQ SEQUENCE 89 AA; 10196 MW; F86024AAB173723 CRC64;

Query Match 79.8%; Score 67; DB 2; Length 89;
 Best Local Similarity 81.8%; Pred. No. 0.035;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
 DB 53 HNNHNNHNNH 63

RESULT 37

Q8DWX2 PRELIMINARY; PRT; 105 AA.

AC Q8DWX2; 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DE Hypothetical protein SAG2089.
 GN OrderedLocustNames=SAG2089;
 OS Streptococcus agalactiae (serotype V).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 RN [1] NCBI_TaxID=216466;
 RP SEQUENCE FROM N.A.

RC STRAIN=2603 V/R / Serotype V;
 RX MEDLINE=22222988; PubMed=12200547; DOI=10.1073/pnas.182380799;
 RA Testelin H., Masiqani V., Cieslewicz M.J., Bisen J.A., Peterson S.N.,
 RA Mesela M.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,
 RA Madoff L.C., Wolf A.M., Beaman M.J., Brinkac L.M., Daugherty S.C.,
 RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
 RA Radu D., Fedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,
 RA Carly H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
 RA Jacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
 RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative genomic analysis of an
 RT emerging human pathogen, serotype V *Streptococcus agalactiae*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
 DR EMBL; AE014286; AAN00948.1; -;
 DR TIGR; SAG2089; -;
 DR InterPro; IPR009711; DUF1292.
 DR Pfam; PF06949; DUF1292; 1.
 KW Complete proteome.
 SQ SEQUENCE 105 AA; 12009 MW; 55CBCTC39F0CDD4D CRC64;

Query Match 79.8%; Score 67; DB 2; Length 105;
 Best Local Similarity 81.8%; Pred. No. 0.041;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
 ID 3 HNNHNNHNNH 13

RESULT 38
 O8E2S4 PRELIMINARY; PRT; 105 AA.

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein gbs2043.
 GN Ordered locus names=gbs2043;
 OS *Streptococcus agalactiae* (serotype III).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=216495;
 RN [1] -
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEM316 / Serotype III;
 RX MEDLINE=22242508; PubMed=12354221;
 RA Glaeser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
 RA Msadek T., Zouine M., Couve E., Lailoui L., Poyart C., Tilleu-Cuot P.,
 RA Kunst F.;
 RT "Genome sequence of *Streptococcus agalactiae*, a pathogen causing
 RT invasive neonatal disease.";
 RL Mol. Microbiol. 45:1499-1513(2002).
 DR EMBL; AL76856; CAD47702.1; -;
 DR Sagalistic; gbs2043; -;
 DR InterPro; IPR009711; DUF1292.
 DR Pfam; PF06949; DUF1292; 1.
 KW Complete proteome.
 SQ SEQUENCE 105 AA; 12009 MW; 55CBCTC39F0CDD4D CRC64;

Query Match 79.8%; Score 67; DB 2; Length 105;
 Best Local Similarity 81.8%; Pred. No. 0.041;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
 ID 3 HNNHNNHNNH 13

RESULT 39
 O8T8F3 PRELIMINARY; PRT; 132 AA.
 ID O8T8F3
 AC O8T8F3;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Rhl protein (Fragment).
 GN Name=rhl;
 OS *Plasmodium falciparum*.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OC NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Taylor H.M., Grainger M., Holder A.A.;
 RL Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ430086; CDB23026.1; -;
 FT NON TER 1 1
 FT NON TER 132 132
 SQ SEQUENCE 132 AA; 15872 MW; D7DAAB81B872B21 CRC64;

Query Match 79.8%; Score 67; DB 2; Length 132;
 Best Local Similarity 83.3%; Pred. No. 0.052;
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
 ID 118 HNNHNNHNNH 129

RESULT 40
 O6C3S6 PRELIMINARY; PRT; 821 AA.

DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Similar to CAGL0C01067g *Candida glabrata* IPF 9648.3.
 GN ORFNames=YAL10E32461g;
 OS *Yarrowia lipolytica* CLIB99.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Dipodascaceae; Yarrowia.
 OC NCBI_TaxID=284591;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIB99;
 RG Genolevures;
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babout A., Barbe V.,
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
 RA Boissrame A., Boyer J., Catolico L., Confantolero F., de Daruvar A.,
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantreay F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
 RA Kerrest A., Kozsl R., Lemaire M., Lesur I., Ma L., Muller H.,
 RA Nicard J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
 RA Pellenz S., Potier S., Richard G.F., Strub M.L., Suleau A.,
 RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
 RA Zentou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
 RA Bouchier C., Caudron B., Scarpelli C., Galliard C., Weisenbach J.,
 RA Wincker P., Souciet J.L.;
 RT "Genome evolution in yeasts."
 RL Nature 430:35-44(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIB99;
 RG Genoscope;
 RL Submitted (Jul-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR382131; CAG80290.1; -;
 SQ SEQUENCE 821 AA; 91509 MW; 0EA01977C3D0D015 CRC64;

Query Match 79.8%; Score 67; DB 2; Length 821;
 Best Local Similarity 81.8%; Pred. No. 0.32;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
 ID 1 HNNHNNHNNH 11

Db 511 HSHNHNHNHN 521

RESULT 41

Q7RVU4

AC Q7RVU4; PRELIMINARY; PRT; 898 AA.

DT 01-MAR-2004 (TREMBLrel. 26, Created)

DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)

DE Hypothetical protein B14D6.440.

GN Name=NCU02826.1;

OS Neurospora crassa.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.

OX NCB1_Taxid=5141;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OR74A;

RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,

RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Renman B.,

RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,

RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,

RA Selltremlkoef C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,

RA Kothe G.O., Jedd G., Mewes W., Steben C., Marcotte E., Greenberg D.,

RA Roy A., Foley K., Naylor J., Thoma N., Barrett R., Gnerre S.,

RA Kamal M., Kamyselis M., Mauceli E., Bielke C., Rudd S., Fishman D.,

RA Desouza C.C., Glas L., Orbach M.J., Berglund J., Voelker R.,

RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,

RA Yarden O., Plamann M., Selter S., Dunlap J., Radford A., Atamayo R.,

RA Nativig D.O., Alex L.A., Mannhaupt G., Ebbels D.J., Freitag M.,

RA Paulsen I., Sachs M.S., Landt E.S., Nussbaum C., Birren B.,

RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";

RL Nature 0:0-0(2003).

CC -1- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

CC EMBL; AABX0100082; EAA34547.1; -;

DR GO; GO:0016021; C:Integral to membrane; IEA.

DR InterPro; IPR004837; NACA_Exmemb.

DR Pfam; PF01699; Na_Ca_ex; 2.

DR KW Hypothetical protein.

SO SEQUENCE 898 AA; 99556 MW; 7BCB3C7B8CE95509 CRC64;

QY Query Match 79.8%; Score 67; DB 2; Length 898;

Best Local Similarity 83.3%; Pred. No. 0.35;

Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1 HSHNHNHNHN 12

641 HSHNHNHNHN 522

RESULT 42

Q7RH05

ID Q7RH05; PRELIMINARY; PRT; 1107 AA.

DT 01-MAR-2004 (TREMBLrel. 26, Created)

DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)

DE Hypothetical protein.

GN Name=PF04191;

OS Plasmodium yoelii yoelii.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.

OX NCB1_Taxid=73239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=17XNL;

RX PubMed=12368865; DOI=10.1038/nature01099;

RA Carlton J.M., Angluoi S.V., Suh B.B., Koof T.W., Persea M.,

RA Silva J.C., Ermolaeva W.D., Allen J.B., Selengut J.D., Koo H.L.,

RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,

RA Shalom S.J., van Aken S.E., Riedmuller S.B., Feldlyum T.V.,

RA Che J.K., Quackenbush J., Sedegh M., Shoaib A., Cummings L.M.,

RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,

RA Cunningham D.A., Preiser P.R., Bergman L.W., Valdivia A.B.,

RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,

RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,

RA Carucci D.J.;

RT "Genome sequence and comparative analysis of the model rodent malaria

RT parasite Plasmodium yoelii yoelii.";

RL Nature 419:512-519(2002).

CC -1- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

CC EMBL; AABL01001251; EAA16013.1; -;

DR KW Hypothetical protein.

SO SEQUENCE 1107 AA; 128769 MW; 9FEF804A2F3AD071 CRC64;

QY Query Match 79.8%; Score 67; DB 2; Length 1107;

Best Local Similarity 75.0%; Pred. No. 0.43;

Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 1 HSHNHNHNHN 12

915 HSHNHNHNHN 926

RESULT 43

Q8WRS6

ID Q8WRS6; PRELIMINARY; PRT; 2957 AA.

DT 01-MAR-2003 (TREMBLrel. 20, Created)

DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)

DE Normocyte-binding protein 1.

GN Name=NBPI;

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCB1_Taxid=5833;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Dd2;

RA MEDLINE=21390493; PubMed=11733572;

RA Kayner J.C., Vargas-Serrato E., Huber C.S., Galinski M.R.,

RA Barnwell J.W.;

RT "A Plasmodium falciparum homologue of Plasmodium vivax reticulocyte

RT binding protein (PvRBP1) defines a trypsin-resistant erythrocyte

RL invasion pathway.";

RL J. Exp. Med. 194:1571-1581(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Dd2;

RA Rayner J.C., Barnwell J.W., Galinski M.R., Tran T., Huber C.S.,

RA Corredor V.;

RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF411931; AAL38220.2; -;

DR Prodom; PD001963; Botulinum; 2.

SO SEQUENCE 2957 AA; 356102 MW; 3BE03A234E78F52E CRC64;

QY Query Match 79.8%; Score 67; DB 2; Length 2957;

Best Local Similarity 83.3%; Pred. No. 1.2;

Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 HSHNHNHNHN 12

2827 HSHNHNHNHN 2838

RESULT 44

Q9DINS

ID Q9DINS; PRELIMINARY; PRT; 114 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-MAR-2004 (TEMBLrel. 26, last annotation update)
 DE Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched
 library, clone:11002F04 product:nerve growth factor receptor
 (TNFRSF16) associated protein 1, full insert sequence.
 GN Name=Ngfrap1; (Mouse).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."; Meth. Enzymol. 303:19-44(1999).
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RA The FANTOM Consortium.
 RT "The RIKEN Genome Exploration Research Group Phase I & II Team;
 the RIKEN Genome Exploration Research Group based on functional annotation of
 60,770 full-length cDNAs."; Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes."; Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 Kono H., Akiyama J., Nishi K., Kitsuunai T., Taahiro H., Itoh M.,
 Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,
 Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer."; Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 Atakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 Hanejaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
 Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
 Sogabe Y., Suzuki H., Tagami M., Tagawa Y., Tanigawa F., Tanaka T.,
 Teijima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK003294; BAB2697.1; -;
 DR MGI: MGI1338016; Ngfrap1.
 DR GO: GO:0005829; C:cytosol; IDA.
 DR GO: GO:0005123; F:death receptor binding; IPI.

DR GO: GO:0006917; P:induction of apoptosis; IDA.
 DR GO: GO:0008625; P:induction of apoptosis via death domain rec. .; IPI.
 DR InterPro: IPR007623; BEX.
 DR Pfam: PF04538; BEX; 1.
 KW Receptor.
 SQ SEQUENCE 114 AA; 13342 MW; 119D070B85B5F802 CRC64;
 Query Match 77.4%; Score 65; DB 2; Length 114;
 Best local Similarity 75.0%; Pred. No. 0.081;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 Db 27 NNNNNNNNNHHH 38
 QY 1 HNNNNNNNNNN 12
 DB 27 NNNNNNNNNHHH 38
 RESULT 45
 ID 09CWN9 PRELIMINARY; PRT; 124 AA.
 AC 09CWN9;
 DT 01-JUN-2001 (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, last sequence update)
 DT 01-MAR-2004 (TEMBLrel. 26, last annotation update)
 DE Mus musculus ES cells cDNA, RIKEN full-length enriched library,
 DE clone:2410015K23 product:nerve growth factor receptor (TNFRSF16)
 DE associated protein 1, full insert sequence.
 GN Name=Ngfrap1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."; Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RA The FANTOM Consortium.
 RT "The RIKEN Genome Exploration Research Group Phase I & II Team;
 the RIKEN Genome Exploration Research Group based on functional annotation of
 60,770 full-length cDNAs."; Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes."; Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 Kono H., Akiyama J., Nishi K., Kitsuunai T., Taahiro H., Itoh M.,
 Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,
 Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

"RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer."
 RL Genome Res. 10:11757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Horii F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kaikawa T., Kato H.,
 RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito K., Saito R., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK010500; BAB26986.1; -;
 DR MGD; MGI:1338016; Nftrapl.
 DR GO; GO:0005829; C:cytosol; IDA.
 DR GO; GO:0005123; F:death receptor binding; IPI.
 DR GO; GO:0006917; P:induction of apoptosis; IDA.
 DR GO; GO:0006625; P:induction of apoptosis via death domain rec. . .; IPI.
 DR InterPro; IPR007623; BEX.
 DR Pfam; PF04538; BEX; 1.
 DR Receptor.
 SQ SEQUENCE 124 AA; 14543 MW; EA1A0F987BECDC3 CRC64;
 Query Match 77.4%; Score 65; DB 2; Length 124;
 Best Local Similarity 75.0%; Pred. No. 0.088;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 HNNHNNHNNHNN 12
 Db 37 NNNNNHNNHNNH 48
 RESULT 46
 Q9WT29 PRELIMINARY; PRT; 124 AA.
 ID Q9WT29;
 AC Q9WT29;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Brain expressed X-linked protein 3 (Nerve growth factor receptor
 DE (TNFRSF16) associated protein 1) (P75NTR-associated cell death
 DE executor).
 GN Name=Ngftrap1; Synonyms=Bex3, Nade;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RX SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Poolled organs;
 RL MEDLINE=99172070; PubMed=10072429; DOI=10.1093/hmg/8.4.611;
 RA Brown A.L., Kay G.F.;
 RT "Bex1, a gene with increased expression in parthenogenetic embryos, is
 RT a member of a novel gene family on the mouse X chromosome."
 RL Hum. Mol. Genet. 8:611-619(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mix FVB/N; TISSUE=Mammary tumor;
 RL MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Colling E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenner C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Bueroc K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stalderon M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Murry D.M., Sodergren E.J., Li X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Kozminski M.I., Skalek U., Smalls D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "mouse and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mix FVB/N; TISSUE=Mammary tumor;
 RL Strausberg R.;
 RN Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RX MEDLINE=20298629; PubMed=10764727; DOI=10.1074/jbc.C000140200;
 RA Mukai J., Hachiya T., Shoji-Hoshino S., Kimura M., Nadeano D.,
 RA Suvarito P., Hanaoka T., Li Y., Irie S., Greene L.A., Sato T.A.;
 RT "NAB2, a p75NTR-associated cell death executor, is involved in signal
 RT transduction mediated by the common neurotrophin receptor p75NTR."
 RL J. Biol. Chem. 275:17566-17570(2000).
 DR EMBL; AF097440; AAD24431.1; -;
 DR EMBL; BC027815; AAB27815.1; -;
 DR EMBL; AF187066; AAF75131.1; -;
 DR MGD; MGI:1338016; Nftrapl.
 DR GO; GO:0005829; C:cytosol; IDA.
 DR GO; GO:0005123; F:death receptor binding; IPI.
 DR GO; GO:0006917; P:induction of apoptosis; IDA.
 DR GO; GO:0006625; P:induction of apoptosis via death domain rec. . .; IPI.
 DR InterPro; IPR007623; BEX.
 DR Pfam; PF04538; BEX; 1.
 DR Receptor.
 SQ SEQUENCE 124 AA; 14542 MW; 3CCCD4F05B66FA61 CRC64;
 Query Match 77.4%; Score 65; DB 2; Length 124;
 Best Local Similarity 75.0%; Pred. No. 0.088;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 HNNHNNHNNHNN 12
 Db 37 NNNNNHNNHNNH 48
 RESULT 47
 Q66GY9 PRELIMINARY; PRT; 124 AA.
 ID Q66GY9;
 AC Q66GY9;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE MGC98067 protein.
 GN Name=MGC98067;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=6355;
 RX SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RL MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative."
 RL Dev. Dyn. 225:384-391(2002).
 RN [2]
 RP SEQUENCE FROM N.A.

RC TISSUE=Embryo; DOI=10.1073/pnas.242603899;
 RA PubMed=12477932; DOI=10.1073/pnas.242603899;
 RX Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
 RA Krausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marisina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Saplento M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshibayashi S., Carninci P., Prange C.,
 RA Raba S.A., Loggiano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fehey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzyaniak M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Gerhard D.S.;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC082126; AAB82126.1; -.
 DR InterPro: IPR007623; BEX.
 DR Pfam: PF04538; BEX; 1.
 SQ SEQUENCE 124 AA; 14542 MW; 3CCCD4F05B66F461 CRC64;
 Query Match 77.4%; Score 65; DB 2; Length 124;
 Best Local Similarity 75.0%; Pred. No. 0.088; Indels 0; Gaps 0;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HNNNNNNNNNN 12
 Db 37 NNNNNNNNNNNH 48
 RESULT 48
 ID Q9D082 PRELIMINARY; PRT; 147 AA.
 AC Q9D082;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DE Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched
 DE library, clone:1190011023 product:nerve growth factor receptor
 DE (TNFRSF16) associated protein 1, full insert sequence.
 GN Name=Ngfrap1;
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RA MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RX Carninci P., Hayashizaki Y.;
 RT "High efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RA MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RX RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;

RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Suganari Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subcloning of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Katsunari T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohata E., Watanabe M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Furumishi Y., Furuo M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurahara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK004531; BAB23350.1; -.
 DR MGD: MGI:1338016; Ngfrap1.
 DR GO: GO:0005829; Cytochrome, IDA.
 DR GO: GO:0005123; Fc receptor binding; IPI.
 DR GO: GO:0006917; P.induction of apoptosis; IDA.
 DR InterPro: IPR007623; BEX.
 DR Pfam: PF04538; BEX; 1.
 KW Receptor.
 SQ SEQUENCE 147 AA; 17124 MW; A15AC7E38BD9470B CRC64;
 Query Match 77.4%; Score 65; DB 2; Length 147;
 Best Local Similarity 75.0%; Pred. No. 0.1;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HNNNNNNNNNN 12
 Db 37 NNNNNNNNNNNH 48
 RESULT 49
 ID Q8Y480 PRELIMINARY; PRT; 303 AA.
 AC Q8Y480;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DE Lmo2575 protein.
 DE Lmo2575 protein.
 GN OrderedlocusNames=Lmo2575;
 OS Listeria monocytogenes.

```

OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=EGD-e / Serovar 1/2a;
MDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaeser F., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
BA Baguero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deloux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,
RA Eutican K.-D., Feilh H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerer U., Kretz J., Kuhn M., Kuntz F., Kurapkat G.,
RA Madueno E., Maitouram A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlend J., Cosset P.,
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL591983; CAD00653.1; -.
DR PIR; AG1396; AG1396.
DR ListRef; LMO2575; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008324; F:cation transporter activity; IEA.
DR PFam; PF01545; C:cation transport; IEA.
DR TIGRFAMs; TIGR01297; CDF; 1.
KW Complete proteome.
SQ SEQUENCE 303 AA; 33140 MW; BE32A94EA3FE81FE CRC64;

Query Match 77.4%; Score 65; DB 2; Length 303;
Best Local Similarity 75.0%; Pred. No. 0.22;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNHNN 12
Db 3 HNHDAHGHNNH 14

RESULT 50
Q927R8 PRELIMINARY; PRT; 303 AA.
AC Q927R8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lin2720 protein.
GN OrderedLocustNames=lin2720;
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ClIP 11262 / Serovar 6a;
MDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaeser F., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
BA Baguero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deloux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,
RA Eutican K.-D., Feilh H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerer U., Kretz J., Kuhn M., Kuntz F., Kurapkat G.,
RA Madueno E., Maitouram A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlend J., Cosset P.,
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL591983; CAD00653.1; -.
DR PIR; AB1772; AB1772.
DR ListRef; LIN2720; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008324; F:cation transporter activity; IEA.

```

```

DR GO; GO:0008324; F:cation transport; IEA.
DR PFam; PF01545; C:cation efflux; 1.
DR TIGRFAMs; TIGR01297; CDF; 1.
KW Complete proteome.
SQ SEQUENCE 303 AA; 33165 MW; AAC6D63CC5473422 CRC64;

Query Match 77.4%; Score 65; DB 2; Length 303;
Best Local Similarity 75.0%; Pred. No. 0.22;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNHNN 12
Db 3 HNHDAHGHNNH 14

RESULT 51
Q71WK3 PRELIMINARY; PRT; 303 AA.
AC Q71WK3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cation efflux family protein.
GN OrderedLocustNames=LMO2365.2548;
OS Listeria monocytogenes (serotype 4b / strain F2365).
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=265669;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15115801; DOI=10.1093/nar/gkh562;
RA Nelson K.E., Fouts D.E., Mongodin E.F., Ravel J., DeBoy R.T.,
RA Kolonay J.F., Raeko D.A., Angiolini S.V., Gill S.R., Paulsen I.T.,
RA Peterson J.D., White O., Nelson W.C., Nielsen W.C., Beaman M.J.,
RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,
RA Haft D.H., Selengut J., Van Aken S.E., Khouri H.M., Fedorova N.,
RA Forbester H.A., Tran B., Kathariou S., Wonderling L.D., Uhlrich G.A.,
RA Bayles D.O., Luchansky J.B., Fraser C.M.;
RT "Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen Listeria monocytogenes reveal new insights into the core genome components of this species.";
RL Nucleic Acids Res. 32:2386-2395(2004).
DR EMBL; AE017330; AF05313.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008324; F:cation transporter activity; IEA.
DR GO; GO:0008324; F:cation transport; IEA.
DR InterPro; IPR002524; C:cation efflux.
DR PFam; PF01545; C:cation efflux; 1.
DR TIGRFAMs; TIGR01297; CDF; 1.
KW Complete proteome.
SQ SEQUENCE 303 AA; 33154 MW; 1E37F94EA3FE8748 CRC64;

Query Match 77.4%; Score 65; DB 2; Length 303;
Best Local Similarity 75.0%; Pred. No. 0.22;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNHNN 12
Db 3 HNHDAHGHNNH 14

RESULT 52
Q6CKC9 PRELIMINARY; PRT; 374 AA.
AC Q6CKC9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similarities with sgdl6006219 Saccharomyces cerevisiae YPR015C.
GN ORFNames=KIA0F11682g;
OS Kluyveromyces lactic NRRL Y-1140.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=284590;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
  Lafontaine I., de Montigny J., Marck C., Nevegilise C., Talla E.,
  Goffard N., Frangoul L., Algie M., Anthouard V., Babour A., Barbe V.,
  Barmy S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
  Boissame A., Boyer J., Cattolico L., Confalonieri F., de Daruvar A.,
  Despons L., Fabre B., Fairhead C., Ferry-Dumazet H., Groppi A.,
  Hantaye F., Hennequin C., Janniaux N., Joyet P., Kachouri R.,
  Kerest A., Koszul R., Lemaitre M., Lesur I., Ma L., Muller H.,
  Nicand J.M., Nkololaki M., Ozdas S., Ozer-Kalogeropoulos O.,
  Pellenz S., Potier S., Richard G.F., Strier-Kalogeropoulos O.,
  Swenene D., Tekala F., Wesolowski-Louvel M., Westhof B., Wirth B.,
  Zenitou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
  Bouchier C., Caudron B., Scarpelli C., Galliard C., Weisenbach J.,
  Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RA Genoscope;
RT Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382126; CAG98318.1; -;
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR007087; ZnF_C2H2.
DR Pfam; PF00096; zf-C2H2; 2.
DR ProDom; PD000003; ZnF_C2H2; 1.
DR SMART; SM00355; ZnF_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
SQ SEQUENCE 374 AA; 41817 MW; 2D04EB7B440F4601 CRC64;

Query Match 77.4%; Score 65; DB 2; Length 374;
Best Local Similarity 81.8%; Pred. No. 0.27;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
Db 102 HNNHNNHNNH 112

RESULT 53
Q9U0J3 PRELIMINARY; PRT; 1010 AA.
AC Q9U0J3;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Hypothetical protein PF04600.
GN Name=PF04600;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ON NCBI_TaxID=36329;
RX MEDLINE=22255708; PubMed=12368867; DOI=10.1038/nature01095;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
  Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
  Bucke C.O., Burrows C., Cherevach I., Chillingworth C.,
  Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
  Cronin A., Davies R., Davis P., Dearden F., Doggett J.,
  Felwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
  Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
  Humphray S., Jagsels K., James K.D., Johnson D., Kenyonou A.,
  Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
  Line A., Maddison M., McLean J., Mooney P., Moulé S., Murphy L.,
  Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
  Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,

```

```

RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
  Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
  RA Taylor J.B., Craig A., Newbold C., Barrell B.G;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).
[2]
RP SEQUENCE FROM N.A.
RC Devlin K., Pain A., Berriman B., Hall N., Bowman S., Churcher C.,
  Harris B., Harris D., Lawson D., Quail M., Barrell B.;
RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035475; CAB62870.2; -;
SQ SEQUENCE 1010 AA; 120658 MW; 90326719C639FFCD CRC64;

Query Match 77.4%; Score 65; DB 2; Length 1010;
Best Local Similarity 90.9%; Pred. No. 0.72;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNNHNNHNNH 12
Db 488 NNNHNNHNNH 498

RESULT 54
Q81IJ7 PRELIMINARY; PRT; 1283 AA.
AC Q81IJ7;
DT 01-MAR-2003 (TREMblrel. 23, Created)
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Prfl_0176;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ON NCBI_TaxID=36329;
RX MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Pung B., White O., Berriman M., Hyman R.W.,
  Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
  Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
  Chan M.S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
  Pereira M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
  Martin D.M., Fairlamb A.H., Fraunholz M.D., Roos D.S., Ralph S.A.,
  RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
  Venter J.C., Carnucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
  Frazer C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
  falciparum.";
RL Nature 419:498-511(2002).
DR EMBL; AE014838; AAN35760.1; -;
DR InterPro; IPR000345; Cytochrome_B5.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 1283 AA; 153000 MW; 2AA707521CCA94D9 CRC64;

Query Match 77.4%; Score 65; DB 2; Length 1283;
Best Local Similarity 75.0%; Pred. No. 0.91;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
Db 830 NNNHNNHNNH 841

RESULT 55
Q9SIB4 PRELIMINARY; PRT; 244 AA.
AC Q9SIB4;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 05-JUL-2004 (TREMblrel. 27, Last annotation update)
DE Putative homeodomain transcription factor (PRESSED FLOWER) (FRS/WOX3

```


Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNNNNNNNH 12
|:|:|:|:|:|:|

Db 5 HSHNHSNHN 16

RESULT 58

Q945N2

PRELIMINARY; PRT; 329 AA.

AC Q945N2; PRELIMINARY; PRT; 329 AA.

DT 01-DEC-2001 (TEMBLrel. 19, Created)

DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)

DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)

DE At2g43970/F6E13.10.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OC NCBI_TaxID=3702;

OX NCBI_TaxID=3702;

RP SEQUENCE FROM N.A.

RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J., Bowser L., Carninci P., Dale J.W., Goldsmith A.D., Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.,

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF412052; AL06505.1; -

DR InterPro; IPR006630; Lupus_La_dom.

DR Pfam; PF05383; La; 1.

DR SMART; SM00715; La; 1.

SQ SEQUENCE 329 AA; 36534 MW; F9366904595C1C4A CRC64;

Query Match 76.2%; Score 64; DB 2; Length 329;
Best local Similarity 75.0%; Pred. No. 0.31;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HNNNNNNNNH 12
|:|:|:|:|:|:|

Db 258 NNNSHNNNNH 269

RESULT 59

O51561

PRELIMINARY; PRT; 451 AA.

AC O51561; PRELIMINARY; PRT; 451 AA.

DT 01-JUN-1998 (TEMBLrel. 06, Created)

DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)

DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)

DE Conserved hypothetical integral membrane protein.

GN OrderedLocustNames=BB0616;

OS Borrelia burgdorferi (Lyme disease spirochete).

OC Bacteria; Spirochaetes; Spirochaetetes; Spirochaetaceae; Borrelia.

OC NCBI_TaxID=139;

RP SEQUENCE FROM N.A.

RA STRAIN=ATCC 35210 / B31;

RA MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;

RA Fraser C.M., Caflens S., Huang W.M., Sutton G.G., Clayton R.A., Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K., Gaim W.D., Dougherty B.A., Tomb J.-F., Fleischmann R.D., Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J., Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D., Gocevane J.D., Weidman J.F., Uitterback T.R., Matthey L., McDonald L.A., Attlich P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K., Roberts K.M., Hatch B., Smith H.O., Venter J.C., "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi."

RL Nature 390:580-586(1997).

DR EMBL; AE001163; AAC66979.1; -

DR PIR; G70176; G70176.

DR TIGR; BB0616; -

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0008324; P:cation transporter activity; IEA.

DR GO; GO:0006812; P:cation transport; IEA.

DR InterPro; IPR002524; Cation efflux.

DR Pfam; PF01545; Cation efflux; 1.

DR TRFAMs; TIGR01297; GDP; 1.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 451 AA; 52049 MW; 66CD0B79243B146B CRC64;

Query Match 76.2%; Score 64; DB 2; Length 451;
Best local Similarity 66.7%; Pred. No. 0.43;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNNNNNNNH 12
|:|:|:|:|:|:|

Db 253 HDHHDHNNHNN 264

RESULT 60

O80567

PRELIMINARY; PRT; 545 AA.

AC O80567; Q9C5X1; PRELIMINARY; PRT; 545 AA.

DT 01-NOV-1998 (TEMBLrel. 08, Created)

DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)

DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)

DE Expressed protein (Hypothetical protein At2g43970) (VIRF-interacting protein FIP) (At2g43970/F6E13.10).

GN Name=At2g43970;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OC NCBI_TaxID=3702;

OX NCBI_TaxID=3702;

RP SEQUENCE FROM N.A.

RA Rounleay S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D., Somerville C.R., Venter J.C.,

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Town C.D., Kaul S.,

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K., Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J., Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shin P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R., Theologis A.,

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RA Mayda E., Tzifira T., Citovsky V.,

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [5]

RP SEQUENCE FROM N.A.

RA Shim P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J., Bowser L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M., Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.,

RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

RN [6]

RP SEQUENCE FROM N.A.

RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Shin P.,

RA Banh J., Bowser L., Carninci P., Chung M.K., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Palm C.J., Pham P.K., Quach H.U., Sakano H., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yu G., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
 RU Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.

RA Shinn P., Chen H., Cheuk R., Kim C.J., Banh J., Bowser L.,
 RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Iehida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
 RA Palm C.J., Quach H.U., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
 RU Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC004005; AAC23405.2; -
 DR EMBL; AY056238; AAL07087.1; -
 DR EMBL; AF332565; AAK06847.1; -
 DR EMBL; AF367277; AAK56266.1; -
 DR EMBL; AF375410; AAK52994.1; -
 DR EMBL; AY129474; AAM91060.1; -
 DR PIR; T00677; T00677.

DR GO; GO:0005737; Cytoplasm; IEA.
 DR GO; GO:0005634; Cytoplasm; IEA.
 DR GO; GO:0030523; Cytoplasm; IEA.
 DR GO; GO:0003723; P:RNA binding; IEA.
 DR GO; GO:0006396; P:RNA processing; IEA.
 DR GO; GO:0006405; P:RNA-nucleus export; IEA.

DR InterPro; IPR002344; lnpus Ia.
 DR InterPro; IPR006630; lnpus Ia.
 DR Pfam; PF05383; Ia; 1.
 DR SMART; SM00715; Ia; 1.
 DR KW Hypothetical protein.

DR SEQUENCE 545 AA; 60589 MW; EIA933261FEIED80 CRC64;
 Query Match 76.2%; Score 64; DB 2; Length 545;
 Best Local Similarity 75.0%; Pred. No. 0.52;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 12
 Db 474 NNNHNNHNNHNN 485

RESULT 61

ID 08ID23 PRELIMINARY; PRT; 576 AA.

AC 08ID23; 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Mitochondrial carrier protein, putative.

GN Name=Pl3 0359;
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=56329;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
 RA Bertram M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
 RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrett B.,
 RU Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the mitochondrial carrier family.
 DR EMBL; AL84509; CAD52806.1; -
 DR GO; GO:0016020; C:mitochondrial carrier family.
 DR GO; GO:0005743; C:mitochondrial inner membrane; IEA.

DR GO; GO:0005488; P:binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR001993; Mitoch carrier.
 DR InterPro; IPR002067; Mlt carrier.

DR Pfam; PF00153; Mito carri; 3.
 DR PRINTS; PR00926; MITOCARRIER.
 DR PROSITE; PSS0920; SOLCAR; 3.
 DR Transmembrane; Transport.

SO SEQUENCE 576 AA; 67860 MW; 786E2845A6F7BDF2 CRC64;

Query Match 76.2%; Score 64; DB 2; Length 576;
 Best Local Similarity 75.0%; Pred. No. 0.55;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 12
 Db 150 HNNHNNHNNHNN 161

RESULT 62

ID 09SG87 PRELIMINARY; PRT; 684 AA.

AC 09SG87; 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Putative RING zinc finger protein.

GN Name=TM13.11;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,
 RA Bonning C.M., Koo H., Fujii C.Y., Uteback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.,
 RU Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
 DR HSRP; Q9LRB7; IYU.
 DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.

DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0016567; P:protein ubiquitination; IEA.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF00097; zf-C3HC4; 1.
 DR SMART; SM00184; RING; 1.

DR PROSITE; PSS0089; ZF_RING_2; 1.
 DR Metal-binding; Zinc; Zinc-finger.
 DR SEQUENCE 684 AA; 76659 MW; 946203A42A7399B1 CRC64;

Query Match 76.2%; Score 64; DB 2; Length 684;
 Best Local Similarity 66.7%; Pred. No. 0.65;
 Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 12
 Db 533 HNNHNNHNNHNN 544

RESULT 63

ID 086AS2 PRELIMINARY; PRT; 786 AA.

AC 086AS2; 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Similar to Dictyostelium discoideum (Slime mold). Developmental
 protein DGA037.

OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=MX4;

RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
 RA Gloeckner G., Eichinger L., Szatranski K., Pachebat J., Dear P.,
 RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
 RA Tungal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
 RT "Sequence and analysis of chromosome 2 of *Dicystotellium discoideum*.";
 RL Nature 418:79-85(2002).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Baumgart C.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the bZIP family.
 DR EMBL, AC116956; AOS1139.1; -;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro: IPR004827; TF bZIP.
 DR Pfam: PF00170; bZIP_1; 1.
 DR SMART; SM00338; BRLZ; 1.
 DR PROSITE; PS50217; BZIP; 1.
 DR DNA-binding; Nuclear protein.
 SQ SEQUENCE 786 AA; 87541 MW; 11E2CE29DFA442C37 CRC64;

Query Match 76.2%; Score 64; DB 2; Length 786;
 Best Local Similarity 75.0%; Pred. No. 0.75;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHHNNHHNN 12
 ||:|||||:
 DB 22 HNNHHNNHHNN 33

RESULT 64
 081KML PRELIMINARY; PRT; 825 AA.

AC 081KML; 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=PT14.0583;
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22255705; PubMed=12358864; DOI=10.1038/nature01097;
 RA Gardner M.J., Hall N., Fung B., White O., Berriman M., Hyman R.W.,
 RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
 RA Bisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
 RA Chan M.S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,
 RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
 RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
 RA Fraser C.M., Barrall B.;
 RT "Genome sequence of the human malaria parasite *Plasmodium falciparum*.";
 RL Nature 419:498-511(2002).
 DR EMBL; AE014825; AAN37196.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 825 AA; 97395 MW; 8A64C49354A1B5 CRC64;

Query Match 76.2%; Score 64; DB 2; Length 825;
 Best Local Similarity 75.0%; Pred. No. 0.79;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHHNNHHNN 12
 ||:|||||:
 DB 166 HNNHHNNHHNN 177

RESULT 65

HMDH_PHYBL STANDARD; PRT; 1176 AA.
 ID HMDH_PHYBL
 AC Q12639;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE 3-hydroxy-3-methylglutaryl-coenzyme A reductase (EC 1.1.1.34) (HMG-CoA reductase).
 GN Name=HMG;
 OS Phycococcus blakesleeanus.
 OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
 OC Phycococcus.
 OX NCBI_TaxID=4837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL 1555;
 RA Ritz-Albert J., Cerdas-Olmedo E., Corrochano L.M.;
 RT "Genes for the metabolism of 3-hydroxy-3-methylglutaryl coenzyme A in the fungus *Phycococcus*.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 836-940 FROM N.A.
 RC STRAIN=NRRL 1555;
 RA Corrochano L.M., Avalos J.;
 RT "Cloning a segment of the gene encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase in *Phycococcus blakesleeanus* and *Gibberella fujikuroi*.";
 RL Exp. Mycol. 16:167-171(1992).
 CC -1- FUNCTION: Involved in the control of cholesterol biosynthesis. It is the rate-limiting enzyme of the sterol biosynthesis.
 CC -1- CATALYTIC ACTIVITY: (R)-mevalonate + CoA + 2 NADPH(+) = (S)-3-hydroxy-3-methylglutaryl-CoA + 2 NADPH.
 CC -1- PATHWAY: Cholesterol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic reticulum.
 CC -1- SIMILARITY: Belongs to the HMG-CoA reductase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X58371; CAB97179.1; -;
 DR PIR: S17345; S17345.
 DR HSSP; P04035; 1HW1.
 DR InterPro; IPR002202; HMG-CoA red.
 DR InterPro; IPR009023; HMG CoA NAD bind.
 DR InterPro; IPR004554; HMG CoA R NADP.
 DR InterPro; IPR009029; HMG CoA sub_bind.
 DR InterPro; IPR00731; SSD_5TM.
 DR Pfam; PF00368; HMG-CoA red; 1.
 DR PRINTS; PR00071; HMGCoARDPASE.
 DR TIGRPFAM; TIGR00533; HMG CoA R NADP; 1.
 DR PROSITE; PS00066; HMG CoA REDUCTASE_1; 1.
 DR PROSITE; PS00318; HMG CoA REDUCTASE_2; 1.
 DR PROSITE; PS01192; HMG CoA REDUCTASE_3; 1.
 DR PROSITE; PS50065; HMG CoA REDUCTASE_4; 1.
 DR PROSITE; PS50156; SSD_1.
 DR KEGG; Cholesterol biosynthesis; Endoplasmic reticulum; Glycoprotein; NADP; Oxidoreductase; Transmembrane
 KW Oxidoreductase; Transmembrane
 FT DOMAIN 1 646
 FT 647 744 Membrane-bound (By similarity).
 FT DOMAIN 745 1176 Catalytic (By similarity).
 FT TRANSMEM 35 55 Potential.
 FT TRANSMEM 300 320 Potential.
 FT TRANSMEM 331 351 Potential.
 FT TRANSMEM 623 643 Potential.
 FT TRANSMEM 1121 1141 Potential.
 FT ACT_SITE 841 841 Charge relay system (By similarity).
 FT ACT_SITE 972 972 Charge relay system (By similarity).
 FT

FT ACT SITE 1048 1048 Charge relay system (by similarity).
 FT ACT SITE 1146 1146 Proton donor (by similarity).
 FT CARBOHYD 224 224 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 238 238 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 553 553 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 556 556 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 670 670 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 904 904 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1013 1013 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1066 1066 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 1176 AA; 127818 MW; 396F958CA1FE7672 CRC64;

Query Match Best Local Similarity 76.2%; Score 64; DB 1; Length 1176;
 Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
 DB 705 HNNHSHSHSHN 716

RESULT 66
 Q7KWP2 PRELIMINARY; PRT; 1485 AA.
 ID Q7KWP2;
 DT 05-JUN-2004 (TREMBLrel. 27, Created)
 DT 05-JUN-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUN-2004 (TREMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 OS Dictyostelium discoideum (slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
 RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
 RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
 RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
 RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum,";
 RL Nature 418:79-85(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Baumgart C.;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC117076; AAC52053.2; -;
 DR InterPro: IPR001849; PH.
 DR PROSITE: PS50003; PH_DOMAIN; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 1485 AA; 168383 MW; 396F958CA1FE7672 CRC64;

Query Match Best Local Similarity 76.2%; Score 64; DB 2; Length 1485;
 Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HNNHNNHNNHNN 12
 DB 269 HNNHSHSHSHN 280

RESULT 67
 Q95PH5 PRELIMINARY; PRT; 1709 AA.
 ID Q95PH5;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Histidine kinase DhkL.
 GN Name=dhkL;
 OS Dictyostelium discoideum (slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
 OX NCBI_TaxID=44689;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Anjard C., Loomis W.F.;
 RT "The histidine kinases of Dictyostelium,";
 RL (in) Inouye M., Dutta R. (eds.);
 RL HISTIDINE KINASES IN SIGNAL TRANSDUCTION, pp.1-0, Academic press, San
 RL Diego (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Anjard C., Loomis W.F.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 CC - SIMILARITY: Contains 1 histidine kinase domain.
 DR EMBL: AF362373; AAK54092.2; -;
 DR HSSP; P3928; 10XK.
 DR DictyBase; DD80191389; dhkL.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0001561; F:two-component response regulator activity; IEA.
 DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
 DR GO; GO:0007600; P:sensory perception; IEA.
 DR GO; GO:0000160; P:two-component signal transduction system (p...); IEA.
 DR InterPro: IPR003594; AtpBind_Atpase.
 DR InterPro: IPR004358; Bact_sens_pr_C.
 DR InterPro: IPR011006; Chey_like.
 DR InterPro: IPR005467; His_kinase.
 DR InterPro: IPR003661; His_kinA_N.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000014; PAC.
 DR InterPro: IPR001789; Response_reg.
 DR Pfam; PF02518; HATPase_C; 1.
 DR Pfam; PF00512; HisKA; 1.
 DR Pfam; PF00989; PAS; 1.
 DR Pfam; PF00072; Response_reg; 1.
 DR PRINTS; PR00344; BCTRLSENSOR.
 DR Prodom; PD000039; Response_reg; 2.
 DR SMART; SM00387; HATPase_c; 1.
 DR SMART; SM00388; HisKA; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 1.
 DR SMART; SM00448; REC; 2.
 DR TIGRFAMs; TIGR00229; sensory_box; 1.
 DR PROSITE; PS50109; HIS_KIN; 1.
 DR PROSITE; PS50112; PAS; 1.
 DR PROSITE; PS50110; RESPONSE_REGULATORY; 2.
 KW Kinase; Phosphorylation; Sensory transduction.
 SQ SEQUENCE 1709 AA; 192602 MW; FOA91C505D5DE178 CRC64;

Query Match Best Local Similarity 76.2%; Score 64; DB 2; Length 1709;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HNNHNNHNNHNN 12
 DB 1216 HNNHSHSHSHN 1227

RESULT 68
 Q84N72 PRELIMINARY; PRT; 118 AA.
 ID Q84N72;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE LEGYC (Fragment).
 OS Amicia glandulosa.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Amicia.
 OX NCBI_TaxID=105934;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22511774; PubMed=12644657; DOI=10.1104/pp.102.016311;
RA Citerne H.L., Luo D., Pennington R.T., Coen E., Cronk Q.C.;
RT "A phylogenomic investigation of CYCLOIDEA-like TCP genes in the
RL Leguminosae.";
RN Plant Physiol. 131:1042-1053(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Citerne H.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY225843; AAC88045.1; -.
DR InterPro: IPR005333; TCP.
DR Pfam: PF03634; TCP; 1.
FT NON TER 1 1
FT NON TER 118 118
SQ SEQUENCE 118 AA; 13737 MW; DB051227B0A0401B CRC64;

Query Match 75.0%; Score 63; DB 2; Length 118;
Best Local Similarity 90.0%; Pred. No. 0.15;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNN 10
DB 47 NNNNNNNNN 56

RESULT 69
Q6AN47 PRELIMINARY; PRT; 250 AA.
ID Q6AN47
AC Q6AN47;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=DPI498;
OS Desulfotalea psychrophila.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
OC Desulfobulbaceae; Desulfotalea.
OX NCBI_TaxID=84980;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LSV54 / DSM 12343;
RX PubMed=15305914;
RA Rabus R., Ruepp A., Frickey T., Rattei T., Fartmann B., Stark M.,
RA Buerer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
RA Kleink H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT from permanently cold Arctic sediments.";
RL Environ. Microbiol. 6:887-902(2004).
DR EMBL: C8522870; CAG36227.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0017004; P:cytochrome biogenesis; IEA.
DR InterPro: IPR003834; Cytoch TM.
DR InterPro: IPR011541; NiCo_transpt.
DR Pfam: PF02683; Dsbd; 1.
DR Pfam: PF03824; NiCo; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 250 AA; 27067 MW; 8862374C95B7D46E CRC64;

Query Match 75.0%; Score 63; DB 2; Length 250;
Best Local Similarity 66.7%; Pred. No. 0.32;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
DB 106 HNNHHHHNNHD 117

RESULT 70
AC22_TRIRE STANDARD; PRT; 341 AA.
ID AC22_TRIRE

```

```

AC Q66WN6;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Transcription factor ACB11.
GN Name=ace2;
OS Trichoderma reesei (Hypocrea jecorina).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_TaxID=51453;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RC STRAIN=QW9414 / Rut C-30;
RX MEDLINE=21316469; PubMed=11304525; DOI=10.1074/jbc.M003624200;
RA Aro N., Saloelimo A., Ilmen M., Penttilae M.;
RT "ACB11, a novel transcriptional activator involved in regulation of
RT cellulase and xylanase genes of Trichoderma reesei.";
RL J. Biol. Chem. 276:24309-24314(2001).
CC -1- FUNCTION: Positive regulation of the major cellulase and xylanase
CC genes (cbh1, cbh2, eg12 and xyn2). Binds to the 5'-GGTAAATAA-3'
CC site present in the cbh1 promoter.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: Contains 1 Zn(2)-Cys(6) fungal-type binuclear cluster
CC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF220671; AAK69383.1; -.
DR HSP; P07272; IPT1.
DR InterPro: IPR001138; Fungi_TYRCP_N.
DR Pfam: PF00172; Zn_Clnb; 1.
DR PRINTS: PR00054; FUNGALZNCYS.
DR SMART: SM00066; GAL4_1.
DR PROSITE: PS00463; ZN2_CYS_FUNGAL_1; 1.
DR PROSITE: PS50048; ZN2_CYS_FUNGAL_2; 1.
KM Activator; DNA-binding; Metal-binding; Nuclear protein;
KM Transcription regulation; zinc.
FT DNA BIND 7 36 Zn(2)-Cys(6), fungal-type.
FT DOMAIN 48 65 His-rich
SQ SEQUENCE 341 AA; 37725 MW; 190F68D4571F92F5 CRC64;

Query Match 75.0%; Score 63; DB 1; Length 341;
Best Local Similarity 66.7%; Pred. No. 0.44;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
DB 55 HEHSHSHNNN 66

RESULT 71
Q81MS9 PRELIMINARY; PRT; 341 AA.
ID Q81MS9
AC Q81MS9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG31339-PA.
GN ORFNames=CG31439;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;

```

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.H., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Abri J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brodtier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadiot E., Center A., Chandra I.,
 RA Chertys J.M., Cieslewski S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fodor C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok C., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mileshina N.V., Mobarry C., Morris J., Mosheiff A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paclob J.M.,
 RA Palazzolo M., Pittman G.A., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reiter K., Remington K., Saunders R.D., Scheeler P., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Maasman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodger, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang Q., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*,"
 RA Science 287:2185-2195(2000).
 RL [2]
 RN
 RP
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
 RA Paclob J.M., Park S., Pfeiffer B.D., Richards S., Sodegren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers B.W., Gibbs R.A., Rubin G.M.,
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*
 RT *melanogaster* euchromatic genome sequence,"
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RT a genomic perspective,"
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Mista S., Crosby M.A., Mungall C.U., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review,"
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]

RP SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB003751; AN114054.1; -
 DR FlyBase; FBgn0051439; CG31439.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0008061; F:chitin binding; IEA.
 DR GO; GO:0006030; P:chitin metabolism; IEA.
 DR InterPro; IPR02357; Chitin_bind_Pera.
 DR InterPro; IPR002125; dCMP/cyt_deam.
 DR Pfam; PF01607; CBM_14; 1.
 DR SMART; SM00494; ChBD2; 1.
 DR PROSITE; PS50940; CHIT_BIND_I1.
 DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; UNKNOWN 1.
 SQ SEQUENCE 341 AA; 3862 MW; A935A06377885A15 CRC64;

Query Match 75.0%; Score 63; DB 2; Length 341;
 Best Local Similarity 66.7%; Pred. NO. 0.44;
 Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNHNN 12
 Db 209 HNNHNNHNNHNNH 220

RESULT 72
 ID 0757N8 PRELIMINARY; PRT; 380 AA.
 AC 0757N8;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DE Predicted protein.
 GN Name=NCU04300.1;
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxId=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OR74A;
 RA Galagan J.E., Galvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
 RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Putcell S., Rehman B.,
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
 RA Qui D., Iankiev P., Pedersen D., Nelson M., Washburne M.,
 RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
 RA Koche G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
 RA Roy A., Foley K., Naylor J., Thoman N., Barrett R., Gierre S.,
 RA Kamal M., Kamysvessels M., Mauceli E., Bielke C., Rudd S., Fishman D.,
 RA Krysiofova S., Raennsen C., Metzberg R.L., Perkins D.D., Kroken S.,
 RA Cogoni C., Macino G., Carlsbeide D., Li W., Pratt R.J., Osmani S.A.,
 RA Desouza C.C., Glas L., Orbach M.J., Berglund J., Voelker R.,
 RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freilich M.,
 RA Paulsen I., Sachs M.S., Lander E.S., Nupson C., Birren B.,
 RT "The genome sequence of the filamentous Fungus *Neurospora crassa*,"
 RL Nature 410:0-0(2003).
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AABX01000276; BAA31854.1; -
 SQ SEQUENCE 380 AA; 42025 MW; 4B98359F2D3B50DF CRC64;

Query Match 75.0%; Score 63; DB 2; Length 380;
 Best Local Similarity 75.0%; Pred. NO. 0.49;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNHNN 12
 :|||||:|:|

Db 22 NNNNNNNNNNN 33

RESULT 73

0871H5 PRELIMINARY; PRT; 400 AA.

AC Q871H5; 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein B7H23.030.

GN Name=B7H23.030;
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;

RN [1]
 RA Schulte U., Algen V., Hoheisel J., Brandt P., Farmann B., Holland R.,
 RA Nykatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

RN [2]
 RA SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BX294026; CAD71029.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 400 AA; 43989 MW; 5BA58BA67E40A4D4 CRC64;

Query Match 75.0%; Score 63; DB 2; Length 400;
 Best Local Similarity 75.0%; Pred. No. 0.52;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNNNNNNNHN 12
 Db 42 NNNNNNNNNNN 53

RESULT 74

06LFN6 PRELIMINARY; PRT; 592 AA.

AC Q6LFN6; 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Hypothetical protein.
 GN ORFNames=MA6P1.33, PPF0145W;
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;

RN [1]
 RA SEQUENCE FROM N.A.

RA PubMed=12368867; DOI=10.1038/nature01095;
 RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
 RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
 RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
 RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
 RA Cronin A., Davies R., Davies P., Dear P., Dearden F., Doggett J.,
 RA Felwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
 RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
 RA Humphray S., Jagsels K., James D., Johnson D., Kerhornou A., Knight A.,
 RA Kontfortov B., Keyes S., Larke N., Lawson D., Lennard N., Line A.,
 RA Maddison M., McLean J., Mooney P., Moule S., Murphy L., Oliver K.,
 RA Ormond D., Price C., Quail M.A., Rabinowitsch E., Rajandream M.A.,
 RA Rutter S., Rutherford K.M., Sanders M., Simmonds M., Seeger K.,
 RA Sharp S., Smith R., Squares R., Squares S., Stevens K., Taylor K.,
 RA Tivey A., Unwin L., Whitehead S., Woodward J., Sulston J.E., Craig A.,
 RA Newbold C., Barrell B.G;
 RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
 RL Nature 419:527-531(2002).

RN [2]
 RA SEQUENCE FROM N.A.

RA Cherevach I., Davis P., Goodhead I., Stevens K., Mungall K.,

RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
 RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.G;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR382398; CAG25200.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 592 AA; 69666 MW; A55464A8F01CF9A CRC64;

Query Match 75.0%; Score 63; DB 2; Length 592;
 Best Local Similarity 75.0%; Pred. No. 0.76;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNNNNNNNHN 12
 Db 355 NNNNNNNNNNN 366

RESULT 75

Q811Q8 PRELIMINARY; PRT; 753 AA.

AC Q811Q8; 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein PFD0845W.

GN Name=PFD0845W;
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;

RN [1]
 RA SEQUENCE FROM N.A.

RA MEDLINE=2225708; PubMed=12368867; DOI=10.1038/nature01095;
 RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
 RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
 RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
 RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
 RA Cronin A., Davies R., Davies P., Dear P., Dearden F., Doggett J.,
 RA Felwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
 RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
 RA Humphray S., Jagsels K., James K.D., Johnson D., Kerhornou A.,
 RA Knights A., Kontfortov B., Keyes S., Larke N., Lawson D., Lennard N.,
 RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
 RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
 RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
 RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
 RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
 RA Sulston J.E., Craig A., Newbold C., Barrell B.G;
 RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
 RL Nature 419:527-531(2002).
 DR EMBL; AL035476; CAD49216.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 753 AA; 91152 MW; 477F6DA922EC2539 CRC64;

Query Match 75.0%; Score 63; DB 2; Length 753;
 Best Local Similarity 75.0%; Pred. No. 0.97;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNNNNNNNHN 12
 Db 296 NNNNNNNNNNN 307

RESULT 76

Q869Y0 PRELIMINARY; PRT; 941 AA.

AC Q869Y0; 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Similar to plasmodium falciparum. cell differentiation protein rccl1,
 DE putative.

OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidida; Dictyostelium.
 OX NCBI_TaxID=44689;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gjoechner G., Eichinger L., Szafranski K., Pachbat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RN Nature 418:79-85(2002).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (Mar-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL:AC115305; A052297.1; -
DR InterPro: IPR007216; Rcd1.
DR Pfam: PF04078; Rcd1; 1.
SQ SEQUENCE 941 AA; 107408 MW; E01F97A1CF17CA7 CRC64;

Query Match
Best Local Similarity 75.0%; Score 63; DB 2; Length 941;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNHNN 12
DB 245 NNNHNNHNNHNN 256

RESULT 77
OB1554 PRELIMINARY; PRT; 1088 AA.
ID OB1554
AC OB1554;
DT 01-MAR-2003 (TREMblrel. 23, Created)
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Bromodomain protein, putative.
GN ORName=BFL0635c;
OS Plasmodium falciparum (isolate 3D7).
CC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
CX NCBI_TaxID=36329;
[1]
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2225705; PubMed=1236864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Esen J.A., Rutherford K., Salzberg S.L., Craig A., Kyte S.,
RA Chan M.S., Nene V., Shallow S.J., Suh B., Peterson J., Angioli S.,
RA Partea M., Allen V., Selengut J., Haft D., Mather W.W., Vaidya A.B.,
RA Martin D.M., Fairclamb A.H., Fraunholz M.U., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Garucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Frazer C.M., Barrett B.;
RT "Genome sequence of the human malaria parasite Plasmodium falciparum.";
RN Nature 419:498-511(2002).
[2]
RP SEQUENCE FROM N.A.
RA Hyman R.W., Fung E., Conway A., Kurd O., Mao J., Miranda M.,
RA Nakao B., Kowley D., Yamaki T., Wang F., Davis R.W.;
RL Submitted (Jan-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL:AE014846; AAN36216.1; -
DR HSSP: Q92793; IJSP.
DR InterPro: IPR001487; Bromodomain.
DR Pfam: PF00439; Bromodomain; 1.
DR PRINTS: PR00503; BROMODOMAIN.
DR SMART: SM00297; BROMO; 1.
DR PROSITE: PS00633; BROMODOMAIN_1; UNKNOWN_1.
DR PROSITE: PS50014; BROMODOMAIN_2; 1.
SQ SEQUENCE 1088 AA; 125449 MW; 0056614BE1B511CA CRC64;

```

```

QY 1 HNNHNNHNNHNN 12
DB 135 NNNHNNHNNHNN 146

RESULT 78
Q7S316 PRELIMINARY; PRT; 1097 AA.
ID Q7S316
AC Q7S316;
DT 01-MAR-2004 (TREMblrel. 26, Created)
DT 01-MAR-2004 (TREMblrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=NCU07531.1;
OS Neurospora crassa.
CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
CC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
CX NCBI_TaxID=5141;
[1]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Seltrennkoff C.P., Kinsey J.A., Braun E.L., Zeller A., Schulte D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gagne S.,
RA Kamal M., Kamyaseilis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Kryzofova S., Rasmussen C., Metzgerberg R.L., Perkins D.D., Kroken S.,
RA Desouza C.C., Macino G., Catchside D., Li W., Pratt R.J., Omani S.A.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Nativg D.O., Alex L.A., Mannhaupt G., Ebbold D.J., Fretag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nisbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC preliminary data.
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL:AA80100415; EAA29808.1; -
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0005524; F: ATP binding; IEA.
DR GO: GO:0004008; F: copper-exporting ATPase activity; IEA.
DR GO: GO:0016787; F: hydrolase activity; IEA.
DR GO: GO:0016820; F: hydrolase activity, acting on acid anhydrid. .; IEA.
DR GO: GO:0046872; F: metal ion binding; IEA.
DR GO: GO:0046873; F: metal ion transporter activity; IEA.
DR GO: GO:0008152; F: metal ion transport; IEA.
DR GO: GO:0030001; F: metal ion transport; IEA.
DR InterPro: IPR006403; ATPase-IB1 Cu.
DR InterPro: IPR006416; ATPase-IB1 hv.
DR InterPro: IPR001757; ATPase-IB1 hv.
DR InterPro: IPR005834; Dehalo-like hydro.
DR InterPro: IPR008250; El-E2 ATPase-reg.
DR InterPro: IPR006121; HeavyMe_transpc.
DR InterPro: IPR002088; PPTA.
DR Pfam: PF00102; El-E2 ATPase; 1.
DR Pfam: PF00702; Hydrolase; 1.
DR PRINTS: PR00119; CATAPASE.
DR TIGRPFAMs: TIGR01511; ATPase-IB1 Cu; 1.
DR TIGRPFAMs: TIGR01525; ATPase-IB1 hv; 1.
DR PROSITE: PS00154; ATPASE-IB1 E2; UNKNOWN_1.
DR PROSITE: PS00904; PPTA; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 1097 AA; 117668 MW; 0C027C5DFA7AF45 CRC64;

```

```

Query Match
Best Local Similarity 75.0%; Score 63; DB 2; Length 1088;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 HNNHNNHNNHNN 12
Query Match
Best Local Similarity 66.7%; Score 63; DB 2; Length 1097;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```



```

Db          124 HGNHSHDHVNS 135
|||||:|||||
RESULT 79
O812K9      PRELIMINARY;      PRT; 1348 AA.
ID O812K9;
AC O812K9;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DE 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Diacylglycerol kinase, putative (EC 2.7.1.107).
GN Name=PF11485c;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxId=36329;
[1]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368867; DOI=10.1038/nature01095;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Felwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphrey S., Jagsels K., James K.D., Johnson D., Kethornou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinovitch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulterson J.E., Craig A., Newbold C., Barrell B.G.
RT "Sequence of Plasmodium falciparum chromosome 1, 3-9 and 13."
RL Nature 419:527-531 (2002).
DR EMBL; AL929358; CAD51983.1; -.
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR000756; DAGKA.
DR InterPro; IPR001206; DAGKC.
DR InterPro; IPR002219; DAG_PE-bind.
DR Pfam; PF00130; Cl_1; 1.
DR ProDom; PD005043; DAGKC; 1.
DR SMART; SM00109; Cl_2.
DR SMART; SM00045; DAGKA; 1.
DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 2.
DR KINASE; Transferase.
SQ SEQUENCE 1348 AA; 158973 MW; 7523D6F052DB18FD CRC64;

Query Match          75.0%; Score 63; DB 2; Length 1348;
Best Local Similarity 75.0%; Pred. No. 1.7;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 12
   |||:|||||:
Db          816 HNNHNNHNNHNN 827

RESULT 80
O81K08      PRELIMINARY;      PRT; 1425 AA.
ID O81K08;
AC O81K08;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DE 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PF10_0032;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

```

```

OX NCBI_TaxId=36329;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Bisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
RA Perlea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.
RT "Genome sequence of the human malaria parasite Plasmodium falciparum."
RL Nature 419:498-511 (2002).
DR EMBL; AB014829; AAN35230.1; -.
DR HSSP; P25685; IHDJ.
DR GO; GO:0051082; F:unfolded protein binding; IEA.
DR GO; GO:0006457; P:protein folding; IEA.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ; 1.
DR PRINTS; PR00625; DnaJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS50636; DnaJ_1; UNKNOWN_1.
DR PROSITE; PS50076; DnaJ_2; 1.
DR Hypothetical protein.
SQ SEQUENCE 1425 AA; 170850 MW; 121E7E311DC43A85 CRC64;

Query Match          75.0%; Score 63; DB 2; Length 1425;
Best Local Similarity 75.0%; Pred. No. 1.8;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 12
   |||:|||||:
Db          426 HNNHNNHNNHNN 437

RESULT 81
O81LL7      PRELIMINARY;      PRT; 1621 AA.
ID O81LL7;
AC O81LL7;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DE 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PF14_0226;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxId=36329;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Bisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
RA Perlea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.
RT "Genome sequence of the human malaria parasite Plasmodium falciparum."
RL Nature 419:498-511 (2002).
DR EMBL; AB014819; AAN36839.1; -.
DR InterPro; IPR001646; Speptide repeat.
DR Pfam; PF00805; Pentapeptide; 2.
DR Hypothetical protein.
SQ SEQUENCE 1621 AA; 192605 MW; B4492093FDDAD3A4E CRC64;

Query Match          75.0%; Score 63; DB 2; Length 1621;

```

Best Local Similarity 75.0%; Pred. No. 2.1;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HNNHNNHNNH 12
Db 734 NNNNNNNNNH 745

RESULT 82

Q81ED3

ID Q81ED3 PRELIMINARY; PRT; 1840 AA.

AC Q81ED3; 01-MAR-2003 (TReMBLrel. 23, Created)

DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)

DE Hypothetical protein PF13_0101.

GN Name=PF13_0101;

OS Plasmodium falciparum (isolate 3D7).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

RN NCBI_Taxid=36329;

RP SEQUENCE FROM N.A.
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrett B.,
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

RL EMBL; AL844509; CDS2327.1; --

KM Hypothetical protein.

SQ SEQUENCE 1840 AA; 214845 MW; 40250F03F97D635A CRC64;

Query Match 75.0%; Score 63; DB 2; Length 1840;
Best Local Similarity 75.0%; Pred. No. 2.4;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
Db 425 NNNNNNNNNH 436

RESULT 83

Q81HR5

AC Q81HR5; PRELIMINARY; PRT; 2075 AA.

DT 01-MAR-2003 (TReMBLrel. 23, Created)

DE Hypothetical protein.

GN ORFNames=PF11_0464;

OS Plasmodium falciparum (isolate 3D7).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

RN NCBI_Taxid=36329;

RP SEQUENCE FROM N.A.
RA MEDLINE=2225705; Pubmed=1236864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallow S.J., Suh B., Peterson J., Angiolini S.,
RA Martin D.M., Fairclamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall K.,
RA Venter J.C., Carrucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrett B.,
RA "Genome sequence of the human malaria parasite Plasmodium
RT falciparum";
RL Nature 419:498-511(2002).

EMBL; AE014843; AAN36044.1; --

GO; GO:0005524; F:ATP binding; IEA.

GO; GO:0004672; P:protein kinase activity; IEA.

GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

InterPro; IPR011009; Kinase like.

InterPro; IPR00719; Prot_Kinase.

Pfam; PF00069; Kinase; I.

DR ProDom; PD000001; Prot kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Hypothetical_protein.
SQ SEQUENCE 2075 AA; 246145 MW; 53F39B2100BBA494 CRC64;

Query Match 75.0%; Score 63; DB 2; Length 2075;
Best Local Similarity 75.0%; Pred. No. 2.7;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
Db 480 HNNHNNHNNH 491

RESULT 84

Q6LF57

ID Q6LF57 PRELIMINARY; PRT; 2461 AA.

AC Q6LF57; 05-UTL-2004 (TReMBLrel. 27, Created)

DT 05-UTL-2004 (TReMBLrel. 27, Last sequence update)

DE Hypothetical protein.

GN ORFNames=MAL6P1.227, PPF0965C;

OS Plasmodium falciparum (isolate 3D7).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

RN NCBI_Taxid=36329;

RP SEQUENCE FROM N.A.

RA Pubmed=1236867; DOI=10.1038/nature01095;

RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corto C.,
RA Cronin A., Davies R., Davies P., Dear P., Dearden F., Doggett J.,
RA Felwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphrey S., Jagels K., James D., Johnson D., Kethorou A., Knight A.,
RA Kontofortov B., Keyes S., Larke N., Lawson D., Lennard N., Line A.,
RA Maddison M., McLean J., Mooney P., Moule S., Murphy L., Oliver K.,
RA Ormond D., Price C., Quail M.A., Rabinowitsch E., Rajandream M.A.,
RA Sharp S., Smith R., Squares R., Squares S., Stevens K., Taylor K.,
RA Tiley A., Unwin L., Whitehead S., Woodward J., Sulston J.E., Craig A.,
RA Newbold C., Barrett B.G.,
RA "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).

RN Nature 419:527-531(2002).

RP SEQUENCE FROM N.A.

RA Cherevach I., Davis P., Goodhead I., Stevens K., Mungall K.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrett B.G.,
RA Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.

RL EMBL; CR382400; CAG25014.1; --

KM Hypothetical protein

SQ SEQUENCE 2461 AA; 298843 MW; E675FD4CF8B80DEB CRC64;

Query Match 75.0%; Score 63; DB 2; Length 2461;
Best Local Similarity 75.0%; Pred. No. 3.2;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
Db 465 NNNNNNNNNH 476

RESULT 85

Q812Y8

ID Q812Y8 PRELIMINARY; PRT; 2506 AA.

AC Q812Y8;

DT 01-MAR-2003 (TReMBLrel. 23, Created)

DE Hypothetical protein.

GN ORFNames=PF11_0464;

OS Plasmodium falciparum (isolate 3D7).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

RN NCBI_Taxid=36329;

RP SEQUENCE FROM N.A.

RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrett B.,
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

RL EMBL; AL844509; CDS2327.1; --

KM Hypothetical protein.

SQ SEQUENCE 1840 AA; 214845 MW; 40250F03F97D635A CRC64;

DE Hypothetical protein PF10805w.
GN Name=PF10805w;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=1236867; DOI=10.1038/nature01095;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
Chillingworth T., Christodoulou Z., Clark L., Clark R., Cotton C.,
Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
Fellwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
Humphray S., Jagsels K., James K.D., Johnson D., Kethornou A.,
Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
Seeger K., Sharp S., Smith R., Squares R., Stevens K., Stevens K.,
Taylor K., Tivey S., Unwin L., Whitehead S., Woodward J.,
Sulston J.E., Craig A., Newbold C., Barrell B.G;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13."
RL Nature 419:527-531(2002).
DR EMBL; AL929357; CAD51847.1;
DR GO; GO:0000151; Cubiquitin ligase complex; IEA.
DR GO; GO:0004842; Fubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; Fzinc ion binding; IEA.
DR GO; GO:0016567; F:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_ring.
DR SMART; SM00184; RING; 1.
DR Hypothetical protein.
SQ SEQUENCE 2506 AA; 303749 MW; FA21AB5EA5AB3BAB CRC64;
QY 1 HNNHNNHNNHNN 12
Db 1007 HNNHNNHNNHNN 1018
Query Match 75.0%; Score 63; DB 2; Length 2506;
Best Local Similarity 75.0%; Pred. No. 3.2;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 HNNHNNHNNHNN 12
Db 1007 HNNHNNHNNHNN 1018
RESULT 86
Q815J1 PRELIMINARY; PRT; 2558 AA.
AC Q815J1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PF1075w;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=1236864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
Chan M.S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium falciparum."
RL Nature 419:498-511(2002).
RP SEQUENCE FROM N.A.

RA Hyman R.W., Fung E., Conway A., Kurd O., Mao J., Miranda M.,
Nakao B., Rowley D., Tamaki T., Wang F., Davis R.W.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB014847; AAN36301.1; -;
KW Hypothetical protein.
SQ SEQUENCE 2558 AA; 301910 MW; F0237A9818A336DE CRC64;
QY 1 HNNHNNHNNHNN 12
Db 476 HNNHNNHNNHNN 487
Query Match 75.0%; Score 63; DB 2; Length 2558;
Best Local Similarity 75.0%; Pred. No. 3.3;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 HNNHNNHNNHNN 12
Db 476 HNNHNNHNNHNN 487
RESULT 87
Q81BH6 PRELIMINARY; PRT; 5561 AA.
AC Q81BH6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein PF07_0118.
GN Name=PF07_0118;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Seeger K., Murphy L., Harris D., Berriman M., Pain A., Hall N.,
Quail M., Barrell B.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844506; CAD51024.1; -;
KW Hypothetical protein.
SQ SEQUENCE 5561 AA; 671113 MW; CC21E0D26F733FEE CRC64;
QY 1 HNNHNNHNNHNN 12
Db 2966 HNNHNNHNNHNN 2977
Query Match 75.0%; Score 63; DB 2; Length 5561;
Best Local Similarity 75.0%; Pred. No. 7.2;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 HNNHNNHNNHNN 12
Db 2966 HNNHNNHNNHNN 2977
RESULT 88
Q81K84 PRELIMINARY; PRT; 5890 AA.
AC Q81K84;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PF14_0722;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=1236864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
Chan M.S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium falciparum."
RL Nature 419:498-511(2002).
RP SEQUENCE FROM N.A.

DR EMBL: AB014828; AAN37335.1; -
 DR InterPro: IPR006209; EGF like.
 DR PROSITE: PS01186; EGF_2; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 5890 AA; 699041 MW; CBA1DEAFALC5ACDC CRC64;
 Query Match 75.0%; Score 63; DB 2; Length 5890;
 Best Local Similarity 75.0%; Pred. No. 7.6;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HNNNNNNNNNN 12
 DB 3728 HNNNNNNNNNN 3739
 RESULT 89
 Q9M2K4 PRELIMINARY; PRT; 329 AA.
 AC Q9M2K4
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein F9D24.30 (Ac3g58120) (Pelota-like protein)
 DE (Transcription factor bZIPel)
 GN Name=F9D24.30; Synonyms=Ac3g58120, BZIPel;
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC Eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 [1]
 RA SEQUENCE FROM N.A.
 RA D'Angelo M., Verzi A., Modesto D., Pigazzi M., Valle G., Mewes H.W.,
 RA Lemcke K., Mayer K.F.X., Quecier F., Salanoubat M.,
 RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RA SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RA SEQUENCE FROM N.A.
 RA Kim C.J., Chen H., Cheuk R., Shinn P., Bowser L., Carninci P.,
 RA Dale J.M., Hayashizaki Y., Heuan V.W., Ishida J., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
 RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
 RA Yuan S., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.,
 RA Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RA SEQUENCE FROM N.A.
 RA MEDLINE=2088475; PubMed=12093376;
 RA Haas B.J., Volfovsky N., Town C.D., Trouhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.,
 RA "Full-length messenger RNA sequences greatly improve genome
 RA annotation";
 RA Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
 RN [5]
 RA SEQUENCE FROM N.A.
 RA Brover V., Trouhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RA SEQUENCE FROM N.A.
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinzaki K.,
 RA Ecker J., Theologis A., Davis R.W.,
 RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RA SEQUENCE FROM N.A.
 RA TISSUE=flower and young silique;

RA Tiedemann J.;
 RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the bZIP family.
 DR EMBL: AL137081; CAB68150.1; -
 DR EMBL: BT008728; AAP42741.1; -
 DR EMBL: AY087356; AAM64906.1; -
 DR EMBL: AY093053; AAM13052.1; -
 DR EMBL: AF401300; AAK84223.1; -
 DR PIR: T45972; T45972.
 DR HSSP: P03069; 1YSA.
 DR GO: GO:0005634; C:nucleus; IEA.
 DR GO: GO:0003677; F:DNA binding; IEA.
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro: IPR008917; Euk transcr_DNA.
 DR InterPro: IPR004827; TF_bZIP.
 DR SMART: SM00338; BRLZ_1; bZIP.
 DR PROSITE: PS00217; BZIP_1.
 DR PROSITE: PS00036; BZIP_BASIC; 1.
 KW DNA-binding, Hypothetical protein, Nuclear protein.
 SQ SEQUENCE 329 AA; 36805 MW; 15FDD8B08BDFD39F CRC64;
 Query Match 73.8%; Score 62; DB 2; Length 329;
 Best Local Similarity 66.7%; Pred. No. 0.57;
 Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HNNNNNNNNNN 12
 DB 101 HNNNNNNNNNN 112
 RESULT 90
 Q9YJZ1 PRELIMINARY; PRT; 529 AA.
 AC Q9YJZ1
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein DKFZ564L2123 (Fragment).
 GN Name=DKFZ564L2123;
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 [1]
 RA SEQUENCE FROM N.A.
 RA TISSUE=Brain;
 RA The German cDNA Consortium;
 RA Mandut R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
 RA Fodor G., Han M., Wiemann S.,
 RA Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL050294; CAB43393.1; -
 DR PIR: T08684; T08684.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0046873; F:metal ion transporter activity; IEA.
 DR GO: GO:0030001; F:metal ion transport; IEA.
 DR InterPro: IPR000345; Cyto_heme_BS.
 DR InterPro: IPR007087; ZnF_C2H2.
 DR InterPro: IPR003689; Zn_transp_zfp.
 DR Pfam: PF02535; Zfp_1.
 DR PROSITE: PS00190; CYTOCHROME C_UNKNOW_1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOW_1.
 KW Hypothetical protein.
 FT NON_TER 529
 SQ SEQUENCE 529 AA; 60751 MW; 3A3D595B98DD1C3A CRC64;
 Query Match 73.8%; Score 62; DB 2; Length 529;
 Best Local Similarity 72.7%; Pred. No. 0.92;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HNNNNNNNNNN 11
 DB 124 HSHNHQHSHNH 134

```

RESULT 91
KEN_DROME STANDARD; PRT; 601 AA.
AC 077459;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last annotation update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Probable transcription factor Ken (Ken and Barbie protein).
GN Name=Ken; ORFNames=CG5575;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RX PubMed=10349629; DOI=10.1016/S0925-4773(98)00183-X;
RA Kuehnlein R.P., Chen C.-K., Schuh R.;
RT "A transcription unit at the ken and barbie gene locus encodes a novel
RL Drosophila zinc finger protein.";
RL Mech. Dev. 79:161-164(1998).
RN [2]
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
STAGE.
RX PubMed=14518006; DOI=10.1002/arch.10105;
RA Lukasovich T., Yuge K., Awano W., Aszalos Z., Kondo S., Juni N.,
Yamamoto D.;
RT "The ken and barbie gene encoding a putative transcription factor with
a BTB domain and three zinc finger motifs functions in terminalia
development of Drosophila.";
RL Arch. Insect Biochem. Physiol. 54:77-94(2003).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Ceoliker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abrill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolhakov S.,
Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glaeser K.,
Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris W.,
Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
Houston D., Houston K.A., Howland T.J., Wei M.-H., Ijegaum C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Laeske P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostreil A.,
Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Watsatan D.A., Weinstein G.M., Weisenbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";

```

```

RL Science 287:2185-2195(2000).
RN [4]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Mista S., Crosby M.A., Mungall C.U., Matthews B.B., Campbell K.S.,
Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
Battencourt B.R., Ceoliker S.E., de Grey A.D.N.J., Drysdale R.A.,
Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely; TISSUE=Head;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
George R.A., Guarin H., Krommiller B., Pacle J.M., Park S., Wan K.H.,
Rubin G.M., Ceoliker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
RN [6]
RP FUNCTION: Probable transcription factor, which is required for
terminalia development.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: Expressed from stage 5 in two rather faint
stripes at positions of 64% (anterior domain, AD) and 17%
(posterior domain, PD) egg length. During early gastrulation, at
stage 6, these two stripes become more evident and detectable at
the region posterior to the cephalic furrow and in the hindgut
primordium. The AD disappears as gastrulation proceeds, while the
PD remains. At stage 15, the AD appears again in the foregut, and
PD expression in the hindgut and anal pad. In imaginal disks, it
is ubiquitously expressed in both males and females in genital and
eye-antenna disks. Not expressed in the brain. In genital disks,
it is expressed along the margin of the anterior bulb in males,
while in females it is expressed in the posterior compartment
along the anterior-posterior border, with medial expansion in the
most posterior region.
CC -1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.
Highly expressed in embryos and pupae, and at lower level in
larvae. In adults, it is expressed at higher level in females.
CC -1- MISCELLANEOUS: Defects in ken result in low mating success and
reduced copulation duration. Men and female genitalia often remain
inside the body, and genitalia and analla are missing in some
homozygous flies.
CC -1- SIMILARITY: Contains 1 BTB/POZ domain.
CC -1- SIMILARITY: Contains 3 C2H2-type zinc fingers.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB010260; BAA32683.1; -
CC EMBL; AB010261; BAA32687.1; -
CC EMBL; AB010276; CA01062.1; -
CC EMBL; AE003462; AAF47084.1; -
CC EMBL; AF181647; AAD55433.1; -
CC PIR; T00119; T00119.
CC HSSP; P08048; 7ZNF.
CC Intact; O77459; -.
CC FlyBase; Fggn0011236; ken.
CC GO; GO:0005634; C:nucleus; IMP.
CC GO; GO:0003677; F:DNA binding; IMP.
CC GO; GO:0045497; P:female analla morphogenesis (sensu Holometeta. . .; IMP.
CC GO; GO:0005400; P:female genital morphogenesis; IMP.
CC GO; GO:0007320; P:insemination; IMP.
CC GO; GO:0045496; P:male analla morphogenesis (sensu Holometabola); IMP.

```

DR GO; GO:0030539; P: male genital morphogenesis; IMP.
 DR GO; GO:0006355; P: regulation of transcription, DNA-dependent; IMP.
 DR InterPro; IPR000210; BTB_PoZ.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR007087; ZnF_C2H2.
 DR Pfam; PF00651; BTB; 1.
 DR Pfam; PF00096; ZF_C2H2; 3.
 DR ProDom; PD000003; ZnF_C2H2; 1.
 DR SMART; SM00225; BTB; 1.
 DR SMART; SM00355; ZnF_C2H2; 3.
 DR PROSITE; PS50097; BTB; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 2.
 DR Developmental protein; DNA-binding; Metal-binding; Nuclear protein;
 KW Transcription regulation; Zinc-binding; Zinc-finger.
 FT DOMAIN 33 101
 FT ZN_FING 500 522 C2H2-type 1.
 FT ZN_FING 528 551 C2H2-type 2.
 FT ZN_FING 567 590 C2H2-type 3.
 FT DOMAIN 245 258 His-rich.
 FT DOMAIN 257 278 Asn-rich.
 SO SEQUENCE 601 AA; 67033 MW; C9AAA437DAD48FE8 CRC64;

Query Match 73.8%; Score 62; DB 1; Length 601;
 Best Local Similarity 66.7%; Pred. No. 1;
 Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 1 HNNHNNHNNH 12
 Db 250 HHHHHHHNNNN 261

RESULT 92
 ID YLH3_SCHPO STANDARD; PRT; 732 AA.
 AC Q8HGQ3; P78888;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DE 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Hypothetical transport protein C17D4.03c in chromosome I.
 GN OBFAMES=SPAC17D4.03c;
 OS Schizosaccharomyces pombe (Pisision yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 NX NCBI_TaxID=4896;
 RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown N., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Pissner A.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones W., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Stevens K., Sharp S.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
 RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Weiler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleau V., Motlier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Huret S.M.,
 RA Lucas M., Rochet M., Gallierdin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

RA Shipkovski G.V., Usary D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe."
 RL Nature 415:871-880 (2002).
 RN [2]
 RP SEQUENCE OF 377-732 FROM N.A.
 RC STRAIN=PR745;
 RX MEDLINE=98162722; PubMed=9501991;
 RA Yoshioke S., Kato K., Nakai K., Okayama H., Nojima H.;
 RT "Identification of open reading frames in Schizosaccharomyces pombe
 RT CDNA."
 RL DNA Res. 4:363-369 (1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the cation diffusion facilitator (CDF)
 CC transporter (TC 2.A.4) family. SLC30A subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ch/announce/>
 CC or send an email to license@ebi.ebi.ch).
 CC -----
 CC EMBL; AL391744; CAC05733.1; -
 CC EMBL; D89239; BA13900.1; -
 CC GeneDB; SPombe; SPAC17D4.03c; -
 CC InterPro; IPR002524; Cation_efflux.
 CC Pfam; PF01545; Cation_efflux; 1.
 CC TIGRFams; TIGR01297; CDF; 1.
 KW Hypothetical protein; Transmembrane; Transport.
 FT TRANSMEM 48 70
 FT TRANSMEM 77 99 Potential.
 FT TRANSMEM 119 141 Potential.
 FT TRANSMEM 162 184 Potential.
 FT TRANSMEM 189 208 Potential.
 FT TRANSMEM 220 242 Potential.
 FT TRANSMEM 265 287 Potential.
 FT TRANSMEM 349 371 Potential.
 FT TRANSMEM 381 400 Potential.
 FT TRANSMEM 421 443 Potential.
 FT TRANSMEM 453 475 Potential.
 FT TRANSMEM 585 607 Potential.
 FT TRANSMEM 612 634 Potential.
 FT DOMAIN 482 487 Poly-His.
 FT DOMAIN 572 578 Poly-His.
 FT DOMAIN 618 618 L -> S (in Ref. 2).
 FT CONFLICT 665 665 N -> H (in Ref. 2).
 FT CONFLICT 668 668 I -> V (in Ref. 2).
 SO SEQUENCE 732 AA; 82739 MW; DBA5452D4329CB0B CRC64;

Query Match 73.8%; Score 62; DB 1; Length 732;
 Best Local Similarity 72.7%; Pred. No. 1.3;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 HNNHNNHNNH 11
 Db 540 HHHHHNNHNDH 550

RESULT 93
 ID 07RML2 PRELIMINARY; PRT; 747 AA.
 AC 07RML2;
 DT 01-MAR-2004 (TEMBLrel. 26, Created)
 DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
 DE 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN Name=PY02167;
 OS Plasmodium yoelii yoelii;
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NX NCBI_TaxID=73239;
 RN [1]
 SO SEQUENCE FROM N.A.

```

RC STRAIN=17XNL;
RA PubMed=12368865; DOI=10.1038/nature01089;
RA Carlton J.M., Angiolini S.V., Suh B.B., Kooij T.W., Pereira M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway K.F., Bigwell S.L.,
RA Shallow S.J., van Aken S.B., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoarbi A., Cummings L.M.,
RA Florens L., Yates F.R. II, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser F.R., Bergman L.W., Valdivia A.B.,
RA van Lin L.H., Jense C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Frazer C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; ABL01000593; EAA21597.1; -.
KW Hypothetical protein.
SQ SEQUENCE 747 AA; 88190 MW; A4D56CA42F46C984 CRC64;

Query Match 73.8%; Score 62; DB 2; Length 747;
Best Local Similarity 81.8%; Pred. No. 1.3;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NHHNNHHNNH 12
DB 283 NHHNNHHNNH 293

RESULT 94
Q68CR5 PRELIMINARY; PRT; 831 AA.
AC Q68CR5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp781L10106.
GN Name=DKFZp781L10106;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon carcinoma;
RG The German cDNA Consortium;
RA Koehler K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobio G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR49813; CAH18673.1; -.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR003689; Zn_transp_Zip.
DR Pfam; PF02535; Zip; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 831 AA; 94119 MW; E19600ED12BC20ED CRC64;

Query Match 73.8%; Score 62; DB 2; Length 831;
Best Local Similarity 72.7%; Pred. No. 1.4;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HHHNNHHNNH 11
DB 124 HSHNHSHNNH 134

RESULT 95
Q9ULF5 PRELIMINARY; PRT; 835 AA.
AC Q9ULF5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```

```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE KIAA1265 protein (Fragment).
GN Name=KIAA1265;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K., Kikuno R., Hirokawa M., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 6:337-345(1999).
DR EMBL; AB033091; BAA86579.1; -.
DR Genew; HGNC:20861; SLC39A10.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0030001; F:metal ion transporter activity; IEA.
DR InterPro; IPR000345; Cytochrome_B5.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR003689; Zn_transp_Zip.
DR Pfam; PF02535; Zip; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN_1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 835 AA; 94559 MW; 851503EFC02F1391 CRC64;

Query Match 73.8%; Score 62; DB 2; Length 835;
Best Local Similarity 72.7%; Pred. No. 1.5;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HHHNNHHNNH 11
DB 128 HSHNHSHNNH 138

RESULT 96
Q86H89 PRELIMINARY; PRT; 911 AA.
AC Q86H89;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gleocier G., Bichinger L., Szefranski K., Pachepat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Grigo R., Kumpf K.,
RA Tungeel B., Cox B., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum."
RT Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116960; AAO53134.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008370; F:zinc ion binding; IEA.
DR InterPro; IPR007087; Znf_C2H2.
DR SMART; SM00355; Znf_C2H2; 4.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
KW Hypothetical protein.

```



```

SQ SEQUENCE 911 AA; 102289 MW; DA816C7DB6435B16 CRC64;
Query Match 73.8%; Score 62; DB 2; Length 911;
Best Local Similarity 72.7%; Pred. No. 1.6;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
DB 117 HGHSHNNHSH 127

RESULT 97
081218 PRELIMINARY; PRT; 1342 AA.
AC 081218;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein PF11590c.
GN Name=PF11590c.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OC NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368867; DOI=10.1038/nature01095;
RA Hall N., Pain A., Bertman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cheruvach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Cotton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagsle K., James K.D., Johnson D., Kerithou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Omond D., Price C., Quail M.A., Rabbittowitch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sultson J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13."
RL Nature 419:557-531(2002).
DR EMBL; AL929359; CAD52004.1; -.
DR GO; GO:0005737; C:Cytoplasm; IEA.
DR GO; GO:0005634; C:Nucleus; IEA.
DR GO; GO:0008262; F:Importin-alpha export receptor activity; IEA.
DR GO; GO:0006915; P:Apoptosis; IEA.
DR GO; GO:0008283; P:Cell proliferation; IEA.
DR InterPro; IPR005043; CAS_CSE1_C.
DR Pfam; PF03378; CAS_CSE1_1.
KW Hypothetical protein.
SQ SEQUENCE 1342 AA; 159113 MW; 8756B1AE30844C31 CRC64;

Query Match 73.8%; Score 62; DB 2; Length 1342;
Best Local Similarity 75.0%; Pred. No. 2.3;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
DB 274 NNNHNNHNNH 285

RESULT 98
SLYD_HAEIN STANDARD; PRT; 190 AA.
AC P44830;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 25-OCT-2004 (Rel. 43, Last annotation update)
DE PKB-type peptidyl-prolyl cis-trans isomerase slyd (EC 5.2.1.8)
OC (PPIase) (Rotamase).
GN Name=slyd; OrderedLocNames=HI0699;

```

```

OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OC NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerevage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G.G., Fitzhugh W., Field C.A., Geacayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodex A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Gnehm C.L., Friedman J.L., Fuhrmann J.L., Georgagen N.S.M.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RD."
RL Science 269:496-512(1995).
CC -!- FUNCTION: PPIases accelerate the folding of proteins (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline
CC (omega=0).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (potential)
CC -!- SIMILARITY: Belongs to the PKB-type PPIase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U32753; AAC2358.1; -.
DR PIR; D64087; D64087.
DR HSP; O52980; 11X5.
DR TIGR; HI0699; -.
DR InterPro; IPR001179; PKBP_PPIase.
DR Pfam; PF00254; PKBP_C; 1.
DR PROSITE; PS50059; PKBP_PPIASE_1.
KW Complete proteome; Isomerase; Metal-binding; Rotamase.
FT DOMAIN 1 95 PPIase, PKB-type.
FT DOMAIN 149 190 Metal-binding (Potential).
FT DOMAIN 141 146 Glu-rich (acidic).
FT DOMAIN 148 190 His-rich (basic).
FT DOMAIN 164 190 Gly-rich.
SQ SEQUENCE 190 AA; 20658 MW; BB7991BA1F1C980D CRC64;

Query Match 72.6%; Score 61; DB 1; Length 190;
Best Local Similarity 72.7%; Pred. No. 0.45;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
DB 172 HNNHNNHNNH 182

RESULT 99
094491 PRELIMINARY; PRT; 271 AA.
AC 094491;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE RING/leucine zipper protein rzip.
GN Name=rzip;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OC NCBI_TaxID=44889;
RN [1]
RP SEQUENCE FROM N.A.

```


RA Kimmel A.R., Ginsburg G.,
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U67917; AAB09553.1; -;
 DR DictyBase; DDB0191162; rzpA.
 DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0016567; P:protein ubiquitination; IEA.
 DR InterPro; IPR001841; Znf_ring.
 DR SMART; SM00184; RING_1.
 DR PROSITE; PS00518; ZF_RING_1; 1.
 DR PROSITE; PS50089; ZF_RING_2; 1.
 KW Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 271 AA; 31329 MW; DAD58879D1EB2931 CRC64;

Query Match 72.6%; Score 61; DB 2; Length 271;
 Best Local Similarity 71.4%; Pred. No. 0.64;
 Matches 10; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

OY 1 HNNHNNHNNH--HN 12
 |||||:|:|
 DB 113 HNNHSHNNNNHNN 126

RESULT 100

O7RCN8 PRELIMINARY; PRT; 607 AA.
 ID O7RCN8
 AC O7RCN8; 01-MAR-2004 (TEMBLrel. 26, Created)
 DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
 DE Hypothetical protein (Fragment).
 GN Name=PY05739;
 OS Plasmodium yoelii yoelii.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 OX NCBI_TaxID=73239;
 OX 1}
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=17XNL;
 RX PubMed=1236865; DOI=10.1038/nature01099;
 RA Carlton J.M., Angluoli S.V., Sun B.B., Kooij T.W., Perteau M.,
 RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
 RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
 RA Shallom S.J., Van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
 RA Cho J.K., Quackenbush J., Sedegah M., Shoaihi A., Cummings L.M.,
 RA Plorens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
 RA Cunningham D.A., Preiser P.R., Bergman L.W., Valdiva A.B.,
 RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
 RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
 RA Carucci D.J.;
 RT "Genome sequence and comparative analysis of the model rodent malaria
 parasite Plasmodium yoelii yoelii.";
 RT Nature 419:512-519(2002).
 RC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL; AABL01001860; EAA17807.1; -;
 DR EMBL; AABL01001860; EAA17807.1; -;
 KW Hypothetical protein.
 KW NON_TER 607
 FT
 SQ SEQUENCE 607 AA; 73453 MW; 8F441BC02F31A77E CRC64;

Query Match 72.6%; Score 61; DB 2; Length 607;
 Best Local Similarity 66.7%; Pred. No. 1.4;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 12
 |||||:|:|
 DB 359 HGHVNNYNDHN 370

Search completed: October 4, 2005, 11:47:02
 Job time : 129 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2005, 11:22:37 ; Search time 60 Seconds
(without alignments)
14,930 Million cell updates/sec

Title: US-09-858-332G-15
Perfect score: 84
Sequence: 1 NHHNNHHNNHHN 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues
Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/prodata/1/aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/aa/PCrUS_COMB.pep:*
6: /cgn2_6/prodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84	100.0	16	1 US-08-346-849-53	Sequence 53, App1
2	84	100.0	16	2 US-08-293-284A-53	Sequence 53, App1
3	84	100.0	16	4 US-08-898-300-53	Sequence 53, App1
4	84	100.0	16	4 US-08-824-513-53	Sequence 53, App1
5	78	92.9	241	4 US-09-248-796A-15369	Sequence 15369, A
6	78	92.9	834	2 US-08-861-664-4	Sequence 4, App1
7	78	92.9	834	2 US-08-396-001-4	Sequence 4, App1
8	78	92.9	834	3 US-09-323-433A-4	Sequence 4, App1
9	78	92.9	834	4 US-09-826-752-4	Sequence 4, App1
10	77	91.7	306	4 US-09-248-796A-21212	Sequence 21212, A
11	76	90.5	601	4 US-09-248-796A-14287	Sequence 14287, A
12	74	88.1	249	4 US-09-248-796A-16005	Sequence 16005, A
13	66	78.6	16	1 US-08-346-849-60	Sequence 60, App1
14	66	78.6	16	1 US-08-346-849-60	Sequence 60, App1
15	66	78.6	16	2 US-08-293-284A-60	Sequence 60, App1
16	66	78.6	16	2 US-08-293-284A-60	Sequence 60, App1
17	66	78.6	16	2 US-08-898-300-60	Sequence 60, App1
18	66	78.6	16	4 US-08-898-300-61	Sequence 61, App1
19	66	78.6	16	4 US-08-824-513-60	Sequence 61, App1
20	66	78.6	16	4 US-08-824-513-61	Sequence 61, App1
21	66	78.6	163	4 US-09-248-796A-22113	Sequence 22113, A
22	66	78.6	447	4 US-09-248-796A-20950	Sequence 20950, A
23	64	76.2	154	4 US-09-248-796A-22802	Sequence 22802, A
24	64	76.2	418	4 US-09-248-796A-18441	Sequence 18441, A
25	63	75.0	109	4 US-09-248-796A-26944	Sequence 26944, A
26	62	73.8	150	4 US-09-563-600A-196	Sequence 196, App
27	59	70.2	175	4 US-09-248-796A-16887	Sequence 16887, A

28	59	70.2	175	4	US-09-248-796A-23457	Sequence 23457, A
29	59	70.2	203	4	US-09-270-767-35326	Sequence 35326, A
30	59	70.2	203	4	US-09-270-767-50543	Sequence 50543, A
31	59	70.2	230	4	US-09-248-796A-21362	Sequence 21362, A
32	59	70.2	392	4	US-09-248-796A-18943	Sequence 18943, A
33	59	70.2	403	4	US-09-248-796A-20669	Sequence 20669, A
34	58	69.0	576	4	US-09-248-796A-20509	Sequence 20509, A
35	57	67.9	116	4	US-09-248-796A-24825	Sequence 24825, A
36	57	67.9	400	4	US-09-543-681A-6151	Sequence 6151, Ap
37	57	67.9	524	4	US-09-248-796A-17786	Sequence 17786, A
38	56	66.7	212	4	US-09-538-092-597	Sequence 597, App
39	56	66.7	355	4	US-09-248-796A-14612	Sequence 14612, A
40	56	66.7	623	4	US-09-949-016-7142	Sequence 7142, Ap
41	55	65.5	381	4	US-09-919-497-96	Sequence 96, App1
42	54.5	64.9	438	4	US-09-248-796A-23919	Sequence 23919, A
43	54	64.3	151	4	US-09-270-767-37338	Sequence 37338, A
44	54	64.3	151	4	US-09-270-767-52555	Sequence 52555, A
45	54	64.3	151	4	US-09-270-767-52555	Sequence 52555, A
46	54	64.3	169	4	US-09-270-767-36080	Sequence 36080, A
47	54	64.3	177	4	US-09-270-767-51297	Sequence 51297, A
48	54	64.3	177	4	US-09-270-767-32436	Sequence 32436, A
49	54	64.3	303	2	US-08-203-532P-2	Sequence 2, App1
50	54	64.3	303	3	US-09-078-465-2	Sequence 2, App1
51	54	64.3	303	3	PCT-US95-01882A-2	Sequence 2, App1
52	54	64.3	342	3	US-09-134-001C-4190	Sequence 4190, Ap
53	54	64.3	363	4	US-09-328-352-4930	Sequence 4930, Ap
54	54	64.3	434	4	US-09-252-991A-30855	Sequence 30855, A
55	54	64.3	434	4	US-09-248-796A-20199	Sequence 20199, A
56	54	64.3	467	4	US-09-657-013-69	Sequence 69, App1
57	54	64.3	467	4	US-09-657-013-70	Sequence 70, App1
58	54	64.3	515	3	US-08-942-012B-32	Sequence 32, App1
59	54	64.3	533	4	US-09-252-991A-23560	Sequence 23560, A
60	54	64.3	620	4	US-09-949-016-6776	Sequence 6776, Ap
61	54	64.3	626	4	US-09-949-016-6776	Sequence 6776, Ap
62	54	64.3	633	3	US-08-557-006C-43	Sequence 43, App1
63	54	64.3	633	4	US-09-538-092-212	Sequence 212, App
64	54	64.3	633	4	US-09-633-328B-3	Sequence 3, App1
65	54	64.3	633	4	US-09-824-735-3	Sequence 3, App1
66	54	64.3	697	4	US-09-949-016-9660	Sequence 9660, Ap
67	54	64.3	763	2	US-08-677-862-2	Sequence 2, App1
68	54	64.3	763	2	US-09-252-571-2	Sequence 2, App1
69	54	64.3	763	3	US-09-434-065-2	Sequence 2, App1
70	54	64.3	763	3	US-08-789-275-4	Sequence 4, App1
71	54	64.3	763	3	US-08-789-275-5	Sequence 5, App1
72	53.5	63.7	254	4	US-09-248-796A-19089	Sequence 19089, A
73	53	63.1	25	4	US-09-721-154-14	Sequence 14, App1
74	53	63.1	60	1	US-08-255-457-1	Sequence 1, App1
75	53	63.1	60	2	US-09-115-032-1	Sequence 1, App1
76	53	63.1	60	2	PCT-US95-05772-1	Sequence 1, App1
77	53	63.1	78	4	US-09-248-796A-22216	Sequence 22216, A
78	53	63.1	244	4	US-09-248-796A-15505	Sequence 15505, A
79	53	63.1	297	4	US-09-489-039A-12802	Sequence 12802, A
80	53	63.1	302	2	US-08-203-532P-4	Sequence 4, App1
81	53	63.1	302	3	US-08-950-860-16	Sequence 16, App1
82	53	63.1	302	3	US-09-078-465-4	Sequence 4, App1
83	53	63.1	302	5	PCT-US95-01882A-4	Sequence 4, App1
84	53	63.1	313	3	US-08-686-528A-3	Sequence 3, App1
85	53	63.1	313	3	US-09-456-287-3	Sequence 3, App1
86	53	63.1	337	3	US-08-686-528A-2	Sequence 2, App1
87	53	63.1	337	3	US-09-456-287-2	Sequence 2, App1
88	53	63.1	337	3	US-09-506-066E-10	Sequence 10, App1
89	53	63.1	414	5	PCT-US92-06840-2	Sequence 2, App1
90	53	63.1	427	4	US-09-506-066E-8	Sequence 8, App1
91	53	63.1	554	4	US-09-949-016-7134	Sequence 7134, Ap
92	53	63.1	576	4	US-09-949-016-8033	Sequence 8033, Ap
93	53	63.1	945	4	US-09-248-796A-15743	Sequence 15743, A
94	53	63.1	1402	4	US-09-248-796A-14503	Sequence 14503, A
95	52.5	62.5	362	4	US-09-248-796A-16633	Sequence 16633, A
96	52.5	62.5	491	4	US-09-248-796A-18483	Sequence 18483, A
97	52	61.9	77	4	US-09-252-991A-19317	Sequence 19317, A
98	52	61.9	79	4	US-09-248-796A-27876	Sequence 27876, A
99	52	61.9	88	4	US-09-270-767-37010	Sequence 37010, A
100	52	61.9	88	4	US-09-270-767-52227	Sequence 52227, A

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/346,849
FILING DATE: 30 NOVEMBER 1994
PRIOR APPLICATION DATA: 07/973,326
FILING DATE: 28 DECEMBER 1992
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008FB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-898-300-53

Query Match 100.0%; Score 84; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.3e-07;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHHHHHHHHN 12
Db 1 HHHHHHHHHN 12

RESULT 4
US-08-824-513-53
Sequence 53, Application US/08824513
Patent No. 6800481

GENERAL INFORMATION:
APPLICANT: Holmes, Todd
APPLICANT: Zhang, Shuangang
APPLICANT: Rich, Alexander
APPLICANT: DiPersio, C. Michael
APPLICANT: Lockshin, Curtis
TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
TITLE OF INVENTION: THEREFOR
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESSES:
ADDRESSER: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173-6799

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/824,513
FILING DATE: March 26, 1997

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/293,284
FILING DATE: August 22, 1998

PRIOR APPLICATION DATA: 07/973,326
FILING DATE: 28 DECEMBER 1992
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240

TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-824-513-53

Query Match 100.0%; Score 84; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.3e-07;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHHHHHHHHN 12
Db 1 HHHHHHHHHN 12

RESULT 5
US-09-248-796A-15369
Sequence 15369, Application US/09248796A
Patent No. 6747137

GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196, 132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 15369
LENGTH: 241
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-15369

Query Match 92.9%; Score 78; DB 4; Length 241;
Best Local Similarity 91.7%; Pred. No. 0.0001;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HHHHHHHHHN 12
Db 16 HHHHHHHHHN 27

RESULT 6
US-08-861-464-4
Sequence 4, Application US/08861464
Patent No. 5874210

GENERAL INFORMATION:
APPLICANT: Guarente, Leonard P.
APPLICANT: Austriaco Jr., Nicanor
APPLICANT: Kennedy, Brian
TITLE OF INVENTION: Genes Determining Cellular Senescence
TITLE OF INVENTION: In Yeast
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSER: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/861,464
FILING DATE: 22-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/396,001
FILING DATE: 28-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09351
FILING DATE: 15-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,408
FILING DATE: 16-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
TELECOMMUNICATION INFORMATION:
TELEPHONE: 781-861-6240
TELEFAX: 781-861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-861-464-4

Query Match 92.9%; Score 78; DB 2; Length 834;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 11
Db 609 HNNHNNHNNH 619

RESULT 7
US-08-396-001-4
Sequence 4, Application US/08396001
Patent No. 5919618
GENERAL INFORMATION:
APPLICANT: Guarente, Leonard P.
APPLICANT: Austriaco Jr., Nicanor
APPLICANT: Claus, James
APPLICANT: Cole, Francesca
APPLICANT: Kennedy, Brian
TITLE OF INVENTION: Genes Determining Cellular Senescence in
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/396,001
FILING DATE: 28-FEB-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-6408A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-396-001-4

Query Match 92.9%; Score 78; DB 2; Length 834;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 11
Db 609 HNNHNNHNNH 619

RESULT 8
US-09-323-433A-4
Sequence 4, Application US/09323433A
Patent No. 6218512
GENERAL INFORMATION:
APPLICANT: Guarente, Leonard P.
APPLICANT: Austriaco Jr., Nicanor
APPLICANT: Claus, James J.
APPLICANT: Cole, Francesca
APPLICANT: Kennedy, Brian
TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESCENCE IN
FILE REFERENCE: 0050.1491-003
CURRENT APPLICATION NUMBER: US/09/323,433A
PRIOR FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: US 08/396,001
PRIOR FILING DATE: 1995-02-28
PRIOR APPLICATION NUMBER: PCT/US94/09351
PRIOR FILING DATE: 1994-08-15
PRIOR APPLICATION NUMBER: US 08/107,408
PRIOR FILING DATE: 1993-08-16
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 834
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-323-433A-4

Query Match 92.9%; Score 78; DB 3; Length 834;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 11
Db 609 HNNHNNHNNH 619

RESULT 9
US-09-826-752-4
Sequence 4, Application US/09826752
Patent No. 6787300
GENERAL INFORMATION:
APPLICANT: Guarente, Leonard P.
APPLICANT: Austriaco Jr., Nicanor
APPLICANT: Claus, James J.
APPLICANT: Cole, Francesca
APPLICANT: Kennedy, Brian
TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESCENCE IN
FILE REFERENCE: 0050.1491-005
CURRENT APPLICATION NUMBER: US/09/826,752
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: US 08/396,001
PRIOR FILING DATE: 1995-02-28
PRIOR APPLICATION NUMBER: PCT/US94/09351
PRIOR FILING DATE: 1994-08-15

PRIOR APPLICATION NUMBER: US 08/107,408
PRIOR FILING DATE: 1993-08-16
PRIOR APPLICATION NUMBER: US 09/323,433
PRIOR FILING DATE: 1999-06-01
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastsEQ for Windows Version 4.0
SEQ ID NO 4
LENGTH: 834
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-826-752-4

Query Match 92.9%; Score 78; DB 4; Length 834;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNH 11
Db 609 HNNNNNNNNH 619

RESULT 10
US-09-248-796A-21212
Sequence 21212, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 21212
LENGTH: 306
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-21212

Query Match 91.7%; Score 77; DB 4; Length 306;
Best Local Similarity 91.7%; Pred. No. 0.00018;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNH 12
Db 160 HNNNNNNNNH 171

RESULT 11
US-09-248-796A-14287
Sequence 14287, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 14287
LENGTH: 601
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-14287

Query Match 90.5%; Score 76; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNNNNNNNNH 12
Db 383 NNNNNNNNNH 393

RESULT 12
US-09-248-796A-16005
Sequence 16005, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 16005
LENGTH: 249
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-16005

Query Match 88.1%; Score 74; DB 4; Length 249;
Best Local Similarity 83.3%; Pred. No. 0.0004;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNH 12
Db 31 HNNNNNNNNH 42

RESULT 13
US-08-346-849-60
Sequence 60, Application US/08346849
Patent No. 5670483
GENERAL INFORMATION:
APPLICANT: Zhang, Shuang
APPLICANT: Lockshin, Curtis
APPLICANT: Rich, Alexander
TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
TITLE OF INVENTION: THEREFOR
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSER: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/346,849
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,326
FILING DATE: 28 DECEMBER 1992

ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-346-849-60

Query Match 78.6%; Score 66; DB 1; Length 16;
Best Local Similarity 81.8%; Pred. No. 0.00035;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 HNNHNNHNNH 11
Db 1 HNHQHNHNNH 11

RESULT 14
US-08-346-849-61
Sequence 61, Application US/08346849
Patent No. 3670483
GENERAL INFORMATION:
APPLICANT: Zhang, Shuguang
APPLICANT: Lockshin, Curtis
APPLICANT: Rich, Alexander
APPLICANT: Holmes, Todd
TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/346,849
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,326
FILING DATE: 28 DECEMBER 1992
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-346-849-61

Query Match 78.6%; Score 66; DB 1; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00035;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 HNNHNNHNNH 12
Db 1 HNHQHNHNNH 12

RESULT 15
US-08-293-284A-60
Sequence 60, Application US/08293284A
Patent No. 5955343
GENERAL INFORMATION:
APPLICANT: Holmes, Todd
APPLICANT: Zhang, Shuguang
APPLICANT: Rich, Alexander
APPLICANT: Dipersio, C. Michael
APPLICANT: Lockshin, Curtis
TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/293,284A
FILING DATE: 22-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,326
FILING DATE: 28-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-293-284A-60

Query Match 78.6%; Score 66; DB 2; Length 16;
Best Local Similarity 81.8%; Pred. No. 0.00035;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
Db 1 HNHQHNHNNH 11

RESULT 16
US-08-293-284A-61
Sequence 61, Application US/08293284A
Patent No. 5955343
GENERAL INFORMATION:
APPLICANT: Holmes, Todd
APPLICANT: Zhang, Shuguang
APPLICANT: Rich, Alexander
APPLICANT: Dipersio, C. Michael

APPLICANT: Lockshin, Curtis
TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
TITLE OF INVENTION: THEREFOR
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/293,284A
FILING DATE: 22-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,326
FILING DATE: 28-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008A
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-293-284A-61

Query Match 78.6%; Score 66; DB 2; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00035;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
1 HNNHNNHNNH 12

Db 1 HNNHNNHNNH 12

RESULT 17
US-08-898-300-60
Sequence 60, Application US/08898300
Patent No. 6548630
GENERAL INFORMATION:
APPLICANT: Zhang, Shuguang
APPLICANT: Lockshin, Curtis
APPLICANT: Rich, Alexander
APPLICANT: Holmes, Todd
TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
TITLE OF INVENTION: THEREFOR
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/898,300
FILING DATE: 22 JULY 1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/346,849
FILING DATE: 30 NOVEMBER 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,326
FILING DATE: 28 DECEMBER 1992
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008FB
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-898-300-60

Query Match 78.6%; Score 66; DB 4; Length 16;
Best Local Similarity 81.8%; Pred. No. 0.00035;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
1 HNNHNNHNNH 11

Db 1 HNNHNNHNNH 11

RESULT 18
US-08-898-300-61
Sequence 61, Application US/08898300
Patent No. 6548630
GENERAL INFORMATION:
APPLICANT: Zhang, Shuguang
APPLICANT: Lockshin, Curtis
APPLICANT: Rich, Alexander
APPLICANT: Holmes, Todd
TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
TITLE OF INVENTION: THEREFOR
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/898,300
FILING DATE: 22 JULY 1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/346,849
FILING DATE: 30 NOVEMBER 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,326
FILING DATE: 28 DECEMBER 1992
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008FB

TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-9240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-898-300-61

Query Match 78.6%; Score 66; DB 4; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00035;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
Db 1 HNNHNNHNNH 12

RESULT 19
US-08-824-513-60
; Sequence 60, Application US/08824513
; Patent No. 6800481
; GENERAL INFORMATION:
; APPLICANT: Holmes, Todd
; APPLICANT: Zhang, Shuangang
; APPLICANT: Rich, Alexander
; APPLICANT: DiPersio, C. Michael
; APPLICANT: Lockshin, Curtis
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,513
; FILING DATE: March 26, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/293,284
; FILING DATE: August 22, 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,326
; FILING DATE: 28 DECEMBER 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: MIT-6008A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 861-6240
; TELEFAX: (781) 861-9540
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-824-513-60

Query Match 78.6%; Score 66; DB 4; Length 16;
Best Local Similarity 81.8%; Pred. No. 0.00035;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 HNNHNNHNNH 11
Db 1 HNNHNNHNNH 11

RESULT 20
US-08-824-513-61
; Sequence 61, Application US/08824513
; Patent No. 6800481
; GENERAL INFORMATION:
; APPLICANT: Holmes, Todd
; APPLICANT: Zhang, Shuangang
; APPLICANT: Rich, Alexander
; APPLICANT: DiPersio, C. Michael
; APPLICANT: Lockshin, Curtis
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,513
; FILING DATE: March 26, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/293,284
; FILING DATE: August 22, 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,326
; FILING DATE: 28 DECEMBER 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: MIT-6008A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 861-6240
; TELEFAX: (781) 861-9540
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-824-513-61

Query Match 78.6%; Score 66; DB 4; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00035;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
Db 1 HNNHNNHNNH 12

RESULT 21
US-09-248-796A-22113
; Sequence 22113, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al

```
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1999-02-12
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 22113
/ LENGTH: 163
/ TYPE: PRT
/ ORGANISM: Candida albicans
/ FEATURE:
/ NAME/KEY: UNSURE
/ LOCATION: (116), (127)
/ OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknc
US-09-248-796A-22113
```

```
Query Match          78.6%; Score 66; DB 4; Length 163;
Best Local Similarity 75.0%; Pred. No. 0.0038;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 HNNNNNNNNNN 12
   |||:|||||
Db 52 HHHHHNPHNHN 63
```

```
RESULT 22
US-09-248-796A-20950
/ Sequence 20950, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR FILING DATE: 1998-08-13
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 20950
/ LENGTH: 447
/ TYPE: PRT
/ ORGANISM: Candida albicans
US-09-248-796A-20950
```

```
Query Match          78.6%; Score 66; DB 4; Length 447;
Best Local Similarity 75.0%; Pred. No. 0.011;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 HNNNNNNNNNN 12
   |||:|||||
Db 61 HHHHHNPHNHN 72
```

```
RESULT 23
US-09-248-796A-22802
/ Sequence 22802, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR FILING DATE: 1998-08-13
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR FILING DATE: 1998-02-13
```

```
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 22802
/ LENGTH: 154
/ TYPE: PRT
/ ORGANISM: Candida albicans
US-09-248-796A-22802
```

```
Query Match          76.2%; Score 64; DB 4; Length 154;
Best Local Similarity 75.0%; Pred. No. 0.007;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 HNNNNNNNNNN 12
   |||:|||||
Db 65 HNNNPHNHNHS 76
```

```
RESULT 24
US-09-248-796A-18441
/ Sequence 18441, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR FILING DATE: 1998-08-13
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 18441
/ LENGTH: 418
/ TYPE: PRT
/ ORGANISM: Candida albicans
/ FEATURE:
/ NAME/KEY: UNSURE
/ LOCATION: (17),(18),(19),(20),(21),(22),(24)
/ OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknc
US-09-248-796A-18441
```

```
Query Match          76.2%; Score 64; DB 4; Length 418;
Best Local Similarity 66.7%; Pred. No. 0.02;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 HNNNNNNNNNN 12
   |||:|||||
Db 25 HNNHHHHHHNHN 36
```

```
RESULT 25
US-09-248-796A-26944
/ Sequence 26944, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR FILING DATE: 1998-08-13
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 26944
/ LENGTH: 109
/ TYPE: PRT
/ ORGANISM: Candida albicans
```

US-09-248-796A-26944

Query Match
Best Local Similarity 75.0%; Score 63; DB 4; Length 109;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:|:
DB 55 HNNHNNHNNH 66

RESULT 26

US-09-663-600A-196
; Sequence 196, Application US/09663600A
; Patent No. 6573068
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: 31.US3.CIP
; CURRENT APPLICATION NUMBER: US/09/663,600A
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; NUMBER OF SEQ ID NOS: 1998-09-04
; SOFTWARE: Patent.pm
; SEQ ID NO 196
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-663-600A-196

Query Match 73.8%; Score 62; DB 4; Length 150;
Best Local Similarity 72.7%; Pred. No. 0.013;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
|:|:|:|:|:|:
DB 124 HSHNHSHNH 134

RESULT 27

US-09-248-796A-16887
; Sequence 16887, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16887
; LENGTH: 175
; TYPE: PRT

ORGANISM: Candida albicans
US-09-248-796A-16887

Query Match 70.2%; Score 59; DB 4; Length 175;
Best Local Similarity 58.3%; Pred. No. 0.043;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:|:
DB 104 HDHHHHHHHH 115

RESULT 28

US-09-248-796A-23457
; Sequence 23457, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 23457
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-23457

Query Match 70.2%; Score 59; DB 4; Length 175;
Best Local Similarity 66.7%; Pred. No. 0.043;
Matches 10; Conservative 1; Mismatches 0; Indels 4; Gaps 1;

QY 2 HNNHNNHNNH 12
|:|:|:|:|:|:
DB 73 HNNHNNHNNH 87

RESULT 29

US-09-270-767-35326
; Sequence 35326, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 35326
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-35326

Query Match 70.2%; Score 59; DB 4; Length 203;
Best Local Similarity 58.3%; Pred. No. 0.05;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:|:
DB 84 HHHHHHHHHH 95

RESULT 30

US-09-270-767-50543
; Sequence 50543, Application US/09270767

```
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 50543
/ LENGTH: 203
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-09-270-767-50543

Query Match          70.2%; Score 59; DB 4; Length 203;
Best Local Similarity 58.3%; Pred. No. 0.05;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1 HNNHNNHNNHNN 12
Db      84 HHHHHHHHHHNN 95

RESULT 31
US-09-248-796A-21362
/ Sequence 21362, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 21362
/ LENGTH: 230
/ TYPE: PRT
/ ORGANISM: Candida albicans
US-09-248-796A-21362

Query Match          70.2%; Score 59; DB 4; Length 230;
Best Local Similarity 58.3%; Pred. No. 0.057;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1 HNNHNNHNNHNN 12
Db      66 HHHHHHHHHHNN 77

RESULT 32
US-09-248-796A-19943
/ Sequence 19943, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 19943
/ LENGTH: 192
```

```
/ TYPE: PRT
/ ORGANISM: Candida albicans
US-09-248-796A-19943

Query Match          70.2%; Score 59; DB 4; Length 392;
Best Local Similarity 81.8%; Pred. No. 0.098;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 NNNHNNHNNHNN 12
Db      96 NNNHNNHNNHNN 106

RESULT 33
US-09-248-796A-20669
/ Sequence 20669, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 20669
/ LENGTH: 403
/ TYPE: PRT
/ ORGANISM: Candida albicans
US-09-248-796A-20669

Query Match          70.2%; Score 59; DB 4; Length 403;
Best Local Similarity 58.3%; Pred. No. 0.1;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1 HNNHNNHNNHNN 12
Db      109 HHHHHHHHHHNS 120

RESULT 34
US-09-248-796A-20509
/ Sequence 20509, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 20509
/ LENGTH: 576
/ TYPE: PRT
/ ORGANISM: Candida albicans
US-09-248-796A-20509

Query Match          69.0%; Score 58; DB 4; Length 576;
Best Local Similarity 66.7%; Pred. No. 0.2;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 HNNHNNHNNHNN 12
Db      535 HNNHNNHNNHNN 546
```

RESULT 35
US-09-248-796A-24825
; Sequence 24825, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 24825
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-24825

Query Match
Best Local Similarity 67.9%; Score 57; DB 4; Length 116;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 12
Db 45 YNNHNNHHYHHH 56

RESULT 36
US-09-543-681A-6151
; Sequence 6151, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6151
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6151

Query Match
Best Local Similarity 67.9%; Score 57; DB 4; Length 400;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 12
Db 51 HHGHHDHHEHNN 62

RESULT 37
US-09-248-796A-17786
; Sequence 17786, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17786
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17786

Query Match
Best Local Similarity 67.9%; Score 57; DB 4; Length 524;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 10
Db 440 NNYNNHNNHNN 449

RESULT 38
US-09-538-092-597
; Sequence 597, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatler Version 0.9
; SEQ ID NO 597
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YML053C
US-09-538-092-597

Query Match
Best Local Similarity 66.7%; Score 56; DB 4; Length 212;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNHNNHNNHNN 12
Db 90 NNNNNNNHNNHNN 101

RESULT 39
US-09-248-796A-14612
; Sequence 14612, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14612

```
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-14612
```

```
Query Match          66.7%; Score 56; DB 4; Length 355;
Best Local Similarity 63.6%; Pred. No. 0.24;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 NNNNNNNNNN 12
Db      143 NHHHHHHHHN 153
```

```
RESULT 40
US-09-949-016-7142
; Sequence 7142, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7142
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7142
```

```
Query Match          66.7%; Score 56; DB 4; Length 623;
Best Local Similarity 63.6%; Pred. No. 0.43;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 NNNNNNNNNN 12
Db      521 HHHHHHHHHN 531
```

```
RESULT 41
US-09-919-497-96
; Sequence 96, Application US/09919497
; Patent No. 6773883
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 96
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (59)..(59)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: UNSURE
; LOCATION: (300)..(300)
; OTHER INFORMATION: Xaa = any amino acid
```

```
; NAME/KEY: UNSURE
; LOCATION: (318)..(318)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: UNSURE
; LOCATION: (330)..(330)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: UNSURE
; LOCATION: (345)..(345)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: UNSURE
; LOCATION: (352)..(352)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: UNSURE
; LOCATION: (367)..(367)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: UNSURE
; LOCATION: (369)..(369)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: UNSURE
; LOCATION: (376)..(376)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: UNSURE
; LOCATION: (378)..(378)
; OTHER INFORMATION: Xaa = any amino acid
US-09-919-497-96
```

```
Query Match          65.5%; Score 55; DB 4; Length 381;
Best Local Similarity 63.6%; Pred. No. 0.37;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 HNNNNNNNNH 11
Db      207 HHHHHHHGHQ 217
```

```
RESULT 42
US-09-248-796A-23919
; Sequence 23919, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 23919
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (3), (4)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-23919
```

```
Query Match          64.9%; Score 54.5; DB 4; Length 438;
Best Local Similarity 75.0%; Pred. No. 0.5;
Matches 9; Conservative 2; Mismatches 0; Indels 1; Gaps 1;
```

```
QY      2 NNNNNNNNHN 12
Db      97 HYNHNNHNNQHN 108
```

```
RESULT 43
US-09-270-767-37338
; Sequence 37338, Application US/09270767
```

```
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ NUMBER OF SEQ ID NOS: 1999-03-17
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 37338
/ LENGTH: 151
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-09-270-767-37338
```

```
Query Match      64.3% Score 54; DB 4; Length 151;
Best Local Similarity 66.7%; Pred. No. 0.2;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNHN 12
Db 5 HNNHNNHNNNN 16
```

```
RESULT 44
US-09-270-767-52555
/ Sequence 52555, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ NUMBER OF SEQ ID NOS: 1999-03-17
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 52555
/ LENGTH: 151
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-09-270-767-52555
```

```
Query Match      64.3% Score 54; DB 4; Length 151;
Best Local Similarity 66.7%; Pred. No. 0.2;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNHN 12
Db 5 HNNHNNHNNNN 16
```

```
RESULT 45
US-09-270-767-36080
/ Sequence 36080, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ NUMBER OF SEQ ID NOS: 1999-03-17
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 36080
/ LENGTH: 169
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
/ FEATURE:
/ OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-36080
```

```
Query Match      64.3% Score 54; DB 4; Length 169;
Best Local Similarity 50.0%; Pred. No. 0.22;
```

```
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNHN 12
Db 112 HSHSHSHSHSHS 123
```

```
RESULT 46
US-09-270-767-51297
/ Sequence 51297, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ NUMBER OF SEQ ID NOS: 1999-03-17
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 51297
/ LENGTH: 169
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
/ FEATURE:
/ OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-51297
```

```
Query Match      64.3% Score 54; DB 4; Length 169;
Best Local Similarity 50.0%; Pred. No. 0.22;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNHN 12
Db 112 HSHSHSHSHSHS 123
```

```
RESULT 47
US-09-270-767-32436
/ Sequence 32436, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ NUMBER OF SEQ ID NOS: 1999-03-17
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 32436
/ LENGTH: 177
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-09-270-767-32436
```

```
Query Match      64.3% Score 54; DB 4; Length 177;
Best Local Similarity 50.0%; Pred. No. 0.23;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNHN 12
Db 110 HSHSHSHSHSHS 121
```

```
RESULT 48
US-09-270-767-47653
/ Sequence 47653, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ NUMBER OF SEQ ID NOS: 1999-03-17
```


; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47653
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47653

Query Match 64.3%; Score 54; DB 4; Length 177;
Best Local Similarity 50.0%; Pred. No. 0.23;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNN 12
|:|:|:|:|:|:
DB 110 HSHSHSHSHS 121

RESULT 49
US-08-203-532F-2

; Sequence 2, Application US/08203532F
; Patent No. 5856121

; GENERAL INFORMATION:
; APPLICANT: Gorski, David H.

; TITLE OF INVENTION: Growth Arrest Homeobox Gene
; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:
ADDRESSEE: Calfee, Halter, and Griswold

; STREET: 800 Superior Avenue
; CITY: Cleveland

; STATE: Ohio
; COUNTRY: U.S.A.

; ZIP: 44114-2688
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; APPLICANT: Walsh, Kenneth
; TITLE OF INVENTION: Growth Arrest Homeobox Gene
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
ADDRESSEE: Calfee, Halter, and Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland

; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2688

; TELEFAX: (216) 241-0816
; TELEPHONE: (216) 622-8200

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

RESULT 51

PCT-US95-01882A-2
; Sequence 2, Application PCT/US9501882A

; GENERAL INFORMATION:
; APPLICANT: Gorski, David H.

; TITLE OF INVENTION: Growth Arrest Homeobox Gene
; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:
ADDRESSEE: Calfee, Halter, and Griswold

; STREET: 800 Superior Avenue
; CITY: Cleveland

; STATE: Ohio
; COUNTRY: U.S.A.

; ZIP: 44114-2688
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

Query Match 64.3%; Score 54; DB 3; Length 303;
Best Local Similarity 50.0%; Pred. No. 0.41;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNN 12
|:|:|:|:|:|:
DB 68 HHHHHHHHHH 79

RESULT 51

PCT-US95-01882A-2
; Sequence 2, Application PCT/US9501882A

; GENERAL INFORMATION:
; APPLICANT: Gorski, David H.

; TITLE OF INVENTION: Growth Arrest Homeobox Gene
; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:
ADDRESSEE: Calfee, Halter, and Griswold

; STREET: 800 Superior Avenue
; CITY: Cleveland

; STATE: Ohio
; COUNTRY: U.S.A.

; ZIP: 44114-2688
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816

REGISTRATION NUMBER: 34829
REFERENCE/DOCKET NUMBER: 22311/00114
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 622-8200
TELEFAX: (216) 241-0816
TELEX: 980499
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 303 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-01882A-2

Query Match 64.3%; Score 54; DB 5; Length 303;
Best Local Similarity 50.0%; Pred. No. 0.41;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNNNHNN 12
|:|:|:|:|:|:
DB 68 HHHHHHHHHH 79

RESULT 52
US-09-134-001C-4190
Sequence 4190, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4190
LENGTH: 342
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4190

Query Match 64.3%; Score 54; DB 3; Length 342;
Best Local Similarity 70.0%; Pred. No. 0.46;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNNNNNNHNN 11
|:|:|:|:|:|:
DB 28 NNNHHHHHSH 37

RESULT 53
US-09-328-352-4930
Sequence 4930, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 4930
LENGTH: 363
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-4930

Query Match 64.3%; Score 54; DB 4; Length 363;

Best Local Similarity 50.0%; Pred. No. 0.49;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNNNHNN 12
|:|:|:|:|:|:
DB 187 HHHHHHHHHH 198

RESULT 54
US-09-252-991A-30855
Sequence 30855, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 30855
LENGTH: 434
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30855

Query Match 64.3%; Score 54; DB 4; Length 434;
Best Local Similarity 50.0%; Pred. No. 0.59;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNNNHNN 12
|:|:|:|:|:|:
DB 289 HSHHHHHHHH 300

RESULT 55
US-09-248-796A-20199
Sequence 20199, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 20199
LENGTH: 434
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-20199

Query Match 64.3%; Score 54; DB 4; Length 434;
Best Local Similarity 63.6%; Pred. No. 0.59;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNNNHNN 11
|:|:|:|:|:|:
DB 289 NNNNNHHHHH 299

RESULT 56
US-09-657-013-69
Sequence 69, Application US/09657013

SEQ ID NO 9643
LENGTH: 620
TYPE: PRT
ORGANISM: Human
US-09-949-016-9643

Query Match
Best Local Similarity 64.3%; Score 54; DB 4; Length 620;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HHHHHHHHHH 12
Db 449 HHHHHHHHHH 460

RESULT 61
US-09-949-016-6776
Sequence 6776, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: PasteSeq for Windows Version 4.0
SEQ ID NO 6776
LENGTH: 626
TYPE: PRT
ORGANISM: Human
US-09-949-016-6776

Query Match
Best Local Similarity 64.3%; Score 54; DB 4; Length 626;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HHHHHHHHHH 12
Db 95 HHHHHHHHHH 106

RESULT 62
US-08-557-006C-43
Sequence 43, Application US/08557006C
Patent No. 6258547
GENERAL INFORMATION:
APPLICANT: Beri, Rajinder K.
APPLICANT: Carling, David A.
APPLICANT: Fordy, Robert A.
TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
FILE REFERENCE: NGAP/PHM37588/UST
CURRENT APPLICATION NUMBER: US/08/557,006C
PRIOR FILING DATE: 1996-03-06
PRIOR APPLICATION NUMBER: PCT/GB94/01093
PRIOR FILING DATE: 1994-05-20
PRIOR APPLICATION NUMBER: GB 9310489.1
PRIOR FILING DATE: 1993-05-21
PRIOR APPLICATION NUMBER: GB 9318010.7
PRIOR FILING DATE: 1993-08-31
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 43
LENGTH: 633
TYPE: PRT
ORGANISM: Yeast

FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)-(633)
OTHER INFORMATION: Yeast SNF1 polypeptide
US-08-557-006C-43

Query Match
Best Local Similarity 64.3%; Score 54; DB 3; Length 633;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HHHHHHHHHH 12
Db 18 HHHHHHHHHH 29

RESULT 63
US-09-538-092-212
Sequence 212, Application US/09538092
Patent No. 6753314
GENERAL INFORMATION:
APPLICANT: Giot, Loic
APPLICANT: Mansfield, Tracy A.
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
FILE REFERENCE: 15966-542
CURRENT APPLICATION NUMBER: US/09/538,092
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/127,352
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/178,965
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 1387
SOFTWARE: CurpaSeqFormatter Version 0.9
SEQ ID NO 212
LENGTH: 633
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
FEATURE:
NAME/KEY: misc feature
LOCATION: (0)-(0)
OTHER INFORMATION: Polypeptide Accession Number YDR477W
US-09-538-092-212

Query Match
Best Local Similarity 64.3%; Score 54; DB 4; Length 633;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HHHHHHHHHH 12
Db 18 HHHHHHHHHH 29

RESULT 64
US-09-633-328B-3
Sequence 3, Application US/09633328B
Patent No. 677587
GENERAL INFORMATION:
APPLICANT: BISARO, DAVID A.
TITLE OF INVENTION: METHOD OF ENHANCING PLANT RESISTANCE TO PATHOGENS
FILE REFERENCE: 22727-04041
CURRENT APPLICATION NUMBER: US/09/633,328B
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: 60/147,613
PRIOR FILING DATE: 1999-08-06
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 3
LENGTH: 633
TYPE: PRT
ORGANISM: Saccharomyces sp.
US-09-633-328B-3

Query Match
Best Local Similarity 64.3%; Score 54; DB 4; Length 633;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:
Db 18 HHHHHHHHHH 29

RESULT 65
US-09-824-735-3
; Sequence 3, Application US/09824735
; Patent No. 6784343
; GENERAL INFORMATION:
; APPLICANT: ZHU, JIAN-KANG
; APPLICANT: LIU, JIPIING
; APPLICANT: ISHITANI, MANABU
; APPLICANT: HALFTER, URSULA
; APPLICANT: KIM, CHEOL-SOO
; TITLE OF INVENTION: PROTEINS AND DNA RELATED TO SALT TOLERANCE IN PLANTS
; FILE REFERENCE: 205645US20
; CURRENT APPLICATION NUMBER: US/09/824,735
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/824,735
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 633
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-824-735-3

Query Match 64.3%; Score 54; DB 4; Length 633;
Best Local Similarity 50.0%; Pred. No. 0.86;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:
Db 18 HHHHHHHHHH 29

RESULT 66
US-09-949-016-9660
; Sequence 9660, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9660
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9660

Query Match 64.3%; Score 54; DB 4; Length 697;
Best Local Similarity 50.0%; Pred. No. 0.95;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:
Db 166 HHHHHHHHHH 177

RESULT 67
US-08-677-862-2
; Sequence 2, Application US/08677862
; Patent No. 5874230
; GENERAL INFORMATION:
; APPLICANT: SONG, Ho Yeong
; APPLICANT: ROTH, Mike
; TITLE OF INVENTION: TRAF2-Associated Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, 34th floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677,862
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brezner, David J
; REGISTRATION NUMBER: 24,774
; REFERENCE/DOCKET NUMBER: T96-005/A63613
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415 781-1989
; TELEFAX: 415 398-3249
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 763 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-677-862-2

Query Match 64.3%; Score 54; DB 2; Length 763;
Best Local Similarity 50.0%; Pred. No. 1;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:~|:|:|:
Db 607 HHHHHHHHHH 618

RESULT 68
US-09-252-571-2
; Sequence 2, Application US/09252571
; Patent No. 5981250
; GENERAL INFORMATION:
; APPLICANT: SONG, Ho Yeong
; APPLICANT: ROTH, Mike
; TITLE OF INVENTION: TRAF2-Associated Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, 34th floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/252,571
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/677,862
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brezner, David J
REGISTRATION NUMBER: 24,774
REFERENCE/DOCKET NUMBER: T96-005/A63613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415 781-1989
TELEFAX: 415 398-3249
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 763 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-252-571-2

Query Match 64.3%; Score 54; DB 2; Length 763;
Best Local Similarity 50.0%; Pred. No. 1;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
|:|:|:|:|:|:
Db 607 HHHHHHHHHH 618

RESULT 69
US-09-434-065-2
Sequence 2, Application US/09434065
Patent No. 6107074
GENERAL INFORMATION:
APPLICANT: SONG, Ho Yeong
APPLICANT: ROTH, Mike
TITLE OF INVENTION: TRAF2-Associated Kinase
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobdach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, 34th floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/434,065
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/677,862
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brezner, David J
REGISTRATION NUMBER: 24,774
REFERENCE/DOCKET NUMBER: T96-005/A63613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415 781-1989
TELEFAX: 415 398-3249
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 763 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-09-434-065-2

Query Match 64.3%; Score 54; DB 3; Length 763;
Best Local Similarity 50.0%; Pred. No. 1;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
|:|:|:|:|:|:
Db 607 HHHHHHHHHH 618

RESULT 70
US-08-789-275-4
Sequence 4, Application US/08789275A
Patent No. 6251664
GENERAL INFORMATION:
APPLICANT: Palleteja, Estivill Xavier
APPLICANT: Pritchard, Melanie
TITLE OF INVENTION: Human Gene Sequence of the Down Syndrome Critical
Patent No. 6251664
TITLE OF INVENTION: Region of Human Chromosome 21, Coding for A Serine-
TITLE OF INVENTION: Theonine Protein Kinase (MNB), Expressed in the
FILE REFERENCE: U 01114-4
CURRENT APPLICATION NUMBER: US/08/789,275A
CURRENT FILING DATE: 1997-01-28
NUMBER OF SEQ ID NOS: 6
SEQ ID NO 4
LENGTH: 763
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
US-08-789-275-4

Query Match 64.3%; Score 54; DB 3; Length 763;
Best Local Similarity 50.0%; Pred. No. 1;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
|:|:|:|:|:|:
Db 607 HHHHHHHHHH 618

RESULT 71
US-08-789-275-5
Sequence 5, Application US/08789275A
Patent No. 6251664
GENERAL INFORMATION:
APPLICANT: Palleteja, Estivill Xavier
APPLICANT: Pritchard, Melanie
TITLE OF INVENTION: Human Gene Sequence of the Down Syndrome Critical
Patent No. 6251664
TITLE OF INVENTION: Region of Human Chromosome 21, Coding for A Serine-
TITLE OF INVENTION: Theonine Protein Kinase (MNB), Expressed in the
FILE REFERENCE: U 01114-4
CURRENT APPLICATION NUMBER: US/08/789,275A
CURRENT FILING DATE: 1997-01-28
NUMBER OF SEQ ID NOS: 6
SEQ ID NO 5
LENGTH: 763
TYPE: PRT
ORGANISM: Rat norvegicus
FEATURE:
US-08-789-275-5

Query Match 64.3%; Score 54; DB 3; Length 763;
Best Local Similarity 50.0%; Pred. No. 1;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12

Db 607 HHHHHHHHHH 618

RESULT 72

US-09-248-796A-19089
; Sequence 19089, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19089
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-19089

Query Match 63.7%; Score 53.5; DB 4; Length 254;

Best Local Similarity 47.4%; Pred. No. 0.4;
Matches 9; Conservative 3; Mismatches 0; Indels 7; Gaps 1;

QY 1 HNNNNNNH-----HNNN 12

Db 111 NNNNNHHFQQQQQQHHN 129

RESULT 73

US-09-721-154-14
; Sequence 14, Application US/09721154
; Patent No. 6651008
; GENERAL INFORMATION:
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Adams, Cynthia
; APPLICANT: Sabry, James
; APPLICANT: Crompton, Anne
; TITLE OF INVENTION: Database system including computer code
; FILE REFERENCE: Cytop007C2
; CURRENT APPLICATION NUMBER: US/09/721,154
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 09/311,596
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Psuedo-sequence
US-09-721-154-14

Query Match 63.1%; Score 53; DB 4; Length 25;

Best Local Similarity 54.5%; Pred. No. 0.044;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNH 11

Db 9 HHHHHHHHHH 19

RESULT 74

US-08-255-457-1
; Sequence 1, Application US/08255457

; Patent No. 5780040

; GENERAL INFORMATION:

; APPLICANT: Plant, Andrew G.

; APPLICANT: Gilbert-Rothstein, Joanne V.

; TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL BINDING

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

STREET: 225 Franklin Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110-2804

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/255,457

FILING DATE:

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Paul C.

REGISTRATION NUMBER: 30,162

REFERENCE/DOCKET NUMBER: 00398/090001

TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 60 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-255-457-1

Query Match 63.1%; Score 53; DB 1; Length 60;

Best Local Similarity 50.0%; Pred. No. 0.11;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNNNNNH 12

Db 11 HHHHHHHHHH 22

RESULT 75

US-09-115-032-1
; Sequence 1, Application US/09115032
; Patent No. 5972348
; GENERAL INFORMATION:
; APPLICANT: Plant, Andrew G.
; APPLICANT: Gilbert-Rothstein, Joanne V.
; APPLICANT: Wright, Andrew
; TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL BINDING
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
ADDRESS: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/115,032
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/255,457
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul C.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-115-032-1

Query Match 63.1%; Score 53; DB 2; Length 60;
Best Local Similarity 50.0%; Pred. No. 0.11;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY 1 HNNHNNHNNH 12
DB 11 HHHHHHTHHH 22

RESULT 76
PCT-US95-05772-1
Sequence 1, Application PC/TUS9505772
GENERAL INFORMATION:
APPLICANT: Plant, Andrew G.
APPLICANT: Gilbert-Rochstein, Joanne V.
TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL
TITLE OF INVENTION: BINDING PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05772
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul C.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-05772-1

Query Match 63.1%; Score 53; DB 5; Length 60;
Best Local Similarity 50.0%; Pred. No. 0.11;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY 1 HNNHNNHNNH 12
DB 11 HHHHHHTHHH 22

RESULT 77
US-09-248-796A-22216
Sequence 22216, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBIC.
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 22216
LENGTH: 78
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-22216

Query Match 63.1%; Score 53; DB 4; Length 78;
Best Local Similarity 54.5%; Pred. No. 0.14;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
QY 1 HNNHNNHNNH 11
DB 68 HHHHHHTHHH 78

RESULT 78
US-09-248-796A-15505
Sequence 15505, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAT
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 15505
LENGTH: 244
TYPE: PRT
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: UNSURE
LOCATION: (4)
US-09-248-796A-15505

Query Match 63.1%; Score 53; DB 4; Length 244;
Best Local Similarity 47.4%; Pred. No. 0.45;
Matches 9; Conservative 1; Mismatches 1; Indels 8; Gaps 1;
QY 1 HNNHNNH-----NNNH 11
DB 186 HNNHNNHGVFLDSGHHH 204

RESULT 79
US-09-489-039A-12802
; Sequence 12802, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12802
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12802

Query Match 63.1%; Score 53; DB 4; Length 297;
Best Local Similarity 50.0%; Pred. No. 0.56;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:
DB 147 HDHHDHHDH 158

RESULT 80
US-08-203-532F-4
; Sequence 4, Application US/08203532F
; Patent No. 5856121
; GENERAL INFORMATION:
; APPLICANT: Gorski, David H.
; APPLICANT: Walsh, Kenneth
; TITLE OF INVENTION: Growth Arrest Homeobox Gene
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter, and Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2688
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,532F
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gollick, Mary E.
; REGISTRATION NUMBER: 34829
; REFERENCE/DOCKET NUMBER: 22311/00114
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816
; TELEX: 980499
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-203-532F-4

Query Match 63.1%; Score 53; DB 2; Length 302;

Best Local Similarity 54.5%; Pred. No. 0.56;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
|:|:|:|:|:
DB 68 HHHHHHHHHH 78

RESULT 81
US-08-950-860-16
; Sequence 16, Application US/08950860B
; Patent No. 6121005
; GENERAL INFORMATION:
; APPLICANT: Fourmiller, Alain
; APPLICANT: Mahfoudi, Abdelrahim
; APPLICANT: Marcieau, Christophe
; APPLICANT: Branellec, Didier
; TITLE OF INVENTION: POLYPEPTIDES COMPRISING DOMAINS OF THE GAX PROTEIN
; TITLE OF INVENTION: IMPLICATED IN THE REPRESSION OF TRANSCRIPTION AND/OR
; TITLE OF INVENTION: INTERACTION WITH OTHER PROTEINS, CORRESPONDING NUCLEIC
; TITLE OF INVENTION: ACIDS, AND THEIR USE
; FILE REFERENCE: 03804.0131
; CURRENT APPLICATION NUMBER: US/08/950,860B
; CURRENT FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 16
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-950-860-16

Query Match 63.1%; Score 53; DB 3; Length 302;
Best Local Similarity 54.5%; Pred. No. 0.56;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
|:|:|:|:|:
DB 68 HHHHHHHHHH 78

RESULT 82
US-09-078-465-4
; Sequence 4, Application US/09078465
; Patent No. 6280969
; GENERAL INFORMATION:
; APPLICANT: Gorski, David H.
; APPLICANT: Walsh, Kenneth
; TITLE OF INVENTION: Growth Arrest Homeobox Gene
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter, and Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2688
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,465
; FILING DATE: 14-MAY-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,532
; FILING DATE: 24-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Gollick, Mary E.
; REGISTRATION NUMBER: 34829
; REFERENCE/DOCKET NUMBER: 22311/00114

TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 622-8200
TELEFAX: (216) 241-0816
TELEX: 980499
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 302 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-078-465-4

Query Match 63.1%; Score 53; DB 3; Length 302;
Best Local Similarity 54.5%; Pred. No. 0.56;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 11
Db 68 HHHHHHHHHH 78

RESULT 83
PCT-US95-01882A-4
Sequence 4, Application PC/TUS9501882A
GENERAL INFORMATION:
APPLICANT: Gorski, David H.
TITLE OF INVENTION: Growth Arrest Homeobox Gene
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calfee, Halter, and Griswold
STREET: 800 Superior Avenue
City: Cleveland
STATE: Ohio
COUNTRY: U.S.A.
ZIP: 44114-2688
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01882A
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Golrick, Mary E.
REGISTRATION NUMBER: 34829
REFERENCE/DOCKET NUMBER: 22311/00114
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 622-8200
TELEFAX: (216) 241-0816
TELEX: 980499
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 302 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-01882A-4

Query Match 63.1%; Score 53; DB 5; Length 302;
Best Local Similarity 54.5%; Pred. No. 0.56;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 11
Db 68 HHHHHHHHHH 78

RESULT 84
US-08-686-528A-3
Sequence 3, Application US/08686528A

Patent No. 6054134
GENERAL INFORMATION:
APPLICANT: LINGWOOD, Clifford A.
TITLE OF INVENTION: HAEMOPHILUS ADHESIN PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis, L.L.P.
STREET: 1737 King Street, Suite 500
City: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22314-2756
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/686,528A
FILING DATE: 26-JUL-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Rea, Teresa Stanek
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 032609-001
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 313 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-686-528A-3

Query Match 63.1%; Score 53; DB 3; Length 313;
Best Local Similarity 50.0%; Pred. No. 0.59;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
Db 111 HDHGDHDDHDD 122

RESULT 85
US-09-456-287-3
Sequence 3, Application US/09456287
Patent No. 6218147
GENERAL INFORMATION:
APPLICANT: LINGWOOD, Clifford A.
TITLE OF INVENTION: HAEMOPHILUS ADHESIN PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis, L.L.P.
STREET: 1737 King Street, Suite 500
City: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22314-2756
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/456,287
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/686,528
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Rea, Teresa Stanek
REGISTRATION NUMBER: 30,427

```

; REFERENCE/DOCKET NUMBER: 032609-001
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-456-287-3

Query Match      63.1%; Score 53; DB 3; Length 313;
Best Local Similarity 50.0%; Pred. No. 0.59;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      1 HNNHNNHNNHNN 12
        |::|::|::|::|
Db      111 HDHKHDHHDH 122

RESULT 86
US-08-686-528A-2
; Sequence 2, Application US/08686528A
; Patent No. 6054134
; GENERAL INFORMATION:
; APPLICANT: LINGWOOD, Clifford A.
; TITLE OF INVENTION: HAEMOPHILUS ADHESIN PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Burns, Doane, Swecker & Mathis, L.L.P.
; STREET: 1737 King Street, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22314-2756
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/686,528A
; FILING DATE: 26-JUL-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Rea, Teresa Stanek
; REGISTRATION NUMBER: 30,427
; REFERENCE/DOCKET NUMBER: 032609-001
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 337 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-686-528A-2

Query Match      63.1%; Score 53; DB 3; Length 337;
Best Local Similarity 50.0%; Pred. No. 0.63;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      1 HNNHNNHNNHNN 12
        |::|::|::|::|
Db      135 HDHKHDHHDH 146

RESULT 87
US-09-456-287-2
; Sequence 2, Application US/09456287
; Patent No. 6218147
; GENERAL INFORMATION:
; APPLICANT: LINGWOOD, Clifford A.
; TITLE OF INVENTION: HAEMOPHILUS ADHESIN PROTEIN
; NUMBER OF SEQUENCES: 4
```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis, L.L.P.
; STREET: 1737 King Street, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22314-2756
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/456,287
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/686,528
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rea, Teresa Stanek
; REGISTRATION NUMBER: 30,427
; REFERENCE/DOCKET NUMBER: 032609-001
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 337 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-456-287-2

Query Match      63.1%; Score 53; DB 3; Length 337;
Best Local Similarity 50.0%; Pred. No. 0.63;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      1 HNNHNNHNNHNN 12
        |::|::|::|::|
Db      135 HDHKHDHHDH 146

RESULT 88
US-09-506-066E-10
; Sequence 10, Application US/09506066E
; Patent No. 6630323
; GENERAL INFORMATION:
; APPLICANT: Scott, Matthew
; APPLICANT: Wharton, Keith
; APPLICANT: Zeng, Wenlin
; TITLE OF INVENTION: Naked Cuticle Genes and their Uses
; FILE REFERENCE: STAN-121
; CURRENT APPLICATION NUMBER: US/09/506,066E
; CURRENT FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/120,646
; PRIOR FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(399)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-09-506-066E-10

Query Match      63.1%; Score 53; DB 4; Length 399;
Best Local Similarity 50.0%; Pred. No. 0.75;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      1 HNNHNNHNNHNN 12
        |::|::|::|::|
```

Db 381 HHHHHHHHHH 392

RESULT 89

PCT-US92-06840-2

Sequence 2, Application PC/TUS9206840

GENERAL INFORMATION:

APPLICANT: Shi, Yang

APPLICANT: Seto, Edward

APPLICANT: Shenk, Thomas

TITLE OF INVENTION: YVI TRANSCRIPTION FACTOR AND METHODS OF

TITLE OF INVENTION: ISOLATING SAME

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESS: Ostrolenk, Faber, Gerb & Soffen

STREET: 1180 Avenue of the Americas - 7th Floor

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-8403

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/06840

FILING DATE: 19920814

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/746,485

FILING DATE: 16-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: Dennis, Manette

REGISTRATION NUMBER: 30,623

REFERENCE/DOCKET NUMBER: M-12594 CIP (1570-8)

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 382-0700

TELEFAX: (212) 382-0888

TELEX: 236925

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 414 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US92-06840-2

Query Match 63.1%; Score 53; DB 5; Length 414;

Best Local Similarity 54.5%; Pred. No. 0.78;

Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHHHHHHHHH 11

Db 70 HHHHHHHHHH 80

RESULT 90

US-09-506-066E-8

Sequence 8, Application US/09506066E

Patent No. 6630323

GENERAL INFORMATION:

APPLICANT: Scott, Matthew

APPLICANT: Wharton, Keith

APPLICANT: Zeng, Wenlin

TITLE OF INVENTION: Naked Cuticle Genes and their Uses

FILE REFERENCE: STAN-121

CURRENT APPLICATION NUMBER: US/09/506,066E

PRIOR APPLICATION NUMBER: 2000-02-17

PRIOR FILING DATE: 1999-02-17

NUMBER OF SEQ ID NOS: 15

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 427
TYPE: PRT
ORGANISM: Mus musculus
US-09-506-066E-8

Query Match 63.1%; Score 53; DB 4; Length 427;
Best Local Similarity 50.0%; Pred. No. 0.81;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 HHHHHHHHHH 12

Db 409 HHHHHHHHHH 420

RESULT 91

US-09-949-016-7134

Sequence 7134, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7134

LENGTH: 556

TYPE: PRT

ORGANISM: Human

US-09-949-016-7134

Query Match 63.1%; Score 53; DB 4; Length 556;
Best Local Similarity 54.5%; Pred. No. 1.1;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHHHHHHHHH 11

Db 176 HHHHHHHHHH 186

RESULT 92

US-09-949-016-8033

Sequence 8033, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 8033

LENGTH: 574

TYPE: PRT

ORGANISM: Human

US-09-949-016-8033

Query Match 63.1%; Score 53; DB 4; Length 574;

Best Local Similarity 54.5%; Pred. No. 1.1;

Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNH 11
|:|:|:|:|:|:
Db 230 HHHHHHHHHH 240

RESULT 93

US-09-248-796A-15743

; Sequence 15743; Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstein et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; PRIOR FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 15743

; LENGTH: 945

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-15743

Query Match 63.1%; Score 53; DB 4; Length 945;
Best Local Similarity 58.3%; Pred. No. 1.8;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;QY 1 HNNNNNNNNH 12
|:|:|:|:|:|:
Db 535 HHHHHNNNNN 546

RESULT 94

US-09-248-796A-14503

; Sequence 14503; Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstein et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; PRIOR FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 14503

; LENGTH: 1402

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-14503

Query Match 63.1%; Score 53; DB 4; Length 1402;
Best Local Similarity 63.6%; Pred. No. 2.7;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;QY 1 HNNNNNNNNH 11
|:|:|:|:|:|:
Db 607 HHHHHNNNNH 617

RESULT 95

US-09-248-796A-16633

; Sequence 16633; Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstein et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; PRIOR FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 16633

; LENGTH: 362

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-16633

Query Match 62.5%; Score 52.5; DB 4; Length 362;
Best Local Similarity 66.7%; Pred. No. 0.8;
Matches 8; Conservative 3; Mismatches 0; Indels 1; Gaps 1;QY 1 HNNNNNNNNH 12
|:|:|:|:|:|:
Db 309 HNH-HHHHHNN 319

RESULT 96

US-09-248-796A-18483

; Sequence 18483; Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstein et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; PRIOR FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 18483

; LENGTH: 491

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-18483

Query Match 62.5%; Score 52.5; DB 4; Length 491;
Best Local Similarity 66.7%; Pred. No. 1.1;
Matches 8; Conservative 3; Mismatches 0; Indels 1; Gaps 1;QY 1 HNNNNNNNNH 12
|:|:|:|:|:|:
Db 388 HNH-HHHHHNN 398

RESULT 97

US-09-252-991A-19317

; Sequence 19317; Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO: 19317
;; LENGTH: 77
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19317

Query Match 61.9%; Score 52; DB 4; Length 77;
Best Local Similarity 54.5%; Pred. No. 0.19;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
|:|:|:|:|
Db 25 HDHDDHDDGH 35

RESULT 98
US-09-248-796A-27876
; Sequence 27876, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Kelch Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 27876
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-27876

Query Match 61.9%; Score 52; DB 4; Length 79;
Best Local Similarity 54.5%; Pred. No. 0.2;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
|:|:|:|:|
Db 60 HHHHHHRRHH 70

RESULT 99
US-09-270-767-37010
; Sequence 37010, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 37010
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-37010

Query Match 61.9%; Score 52; DB 4; Length 88;
Best Local Similarity 54.5%; Pred. No. 0.22;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
|:|:|:|:|
Db 78 HSHTHSHSH 88

RESULT 100
US-09-270-767-52227
; Sequence 52227, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 52227
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-52227

Query Match 61.9%; Score 52; DB 4; Length 88;
Best Local Similarity 54.5%; Pred. No. 0.22;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
|:|:|:|:|
Db 78 HSHTHSHSH 88

Search completed: October 4, 2005, 11:23:51
Job time : 62 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2005, 11:22:37 ; Search time 1241 Seconds
(without alignments)
4.014 Million cell updates/sec

Title: US-09-858-332G-15
Perfect score: 84
Sequence: 1 NHNHNNHNNHNN 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1846076 seqs, 415116000 residues

Total number of hits satisfying chosen parameters: 1846076

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubppaa/US11_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84	100.0	12	9	US-09-858-332-18
2	84	100.0	12	16	US-10-453-410-4
3	84	100.0	12	18	US-10-743-956-33
4	84	100.0	13	14	US-10-055-794-15
5	84	100.0	16	15	US-10-390-472-53
6	84	100.0	55	15	US-10-422-262-16
7	84	100.0	55	15	US-10-448-609-12
8	79	94.0	120	10	US-09-327-750-52
9	78	92.9	834	9	US-09-826-752-4
10	78	92.9	834	17	US-10-912-434-4
11	74	88.1	96	16	US-10-437-963-148051

12	70	83.3	490	17	US-10-732-923-2998	Sequence 2998, Ap
13	69	82.1	94	15	US-10-424-599-200505	Sequence 200505,
14	69	82.1	105	15	US-10-424-599-250721	Sequence 250721,
15	69	82.1	193	16	US-10-437-963-152529	Sequence 152529,
16	69	82.1	1284	14	US-10-296-144-5	Sequence 5, Appl1
17	69	82.1	1284	20	US-11-097-143-42399	Sequence 42399, A
18	68.5	81.5	697	20	US-11-097-147-60	Sequence 21195, A
19	66	78.6	16	15	US-10-390-472-61	Sequence 60, Appl
20	66	78.6	148	16	US-10-425-115-204424	Sequence 61, Appl
21	66	78.6	148	16	US-10-425-115-204424	Sequence 204424,
22	65	77.4	124	10	US-09-327-750-34	Sequence 34, Appl
23	65	77.4	124	10	US-09-327-750-53	Sequence 53, Appl
24	64	76.2	276	15	US-10-424-599-180331	Sequence 180331,
25	64	76.2	451	15	US-10-282-122A-47210	Sequence 47210, A
26	64	76.2	545	16	US-10-739-930-5773	Sequence 5773, Ap
27	64	76.2	1176	15	US-10-041-018-257	Sequence 267, App
28	62	73.8	150	11	US-09-978-360A-708	Sequence 708, App
29	62	73.8	150	14	US-10-319-763-196	Sequence 196, App
30	62	73.8	329	10	US-09-934-455-166	Sequence 166, App
31	62	73.8	329	14	US-10-286-264-24	Sequence 24, Appl
32	62	73.8	329	15	US-10-302-267-38	Sequence 1082, Ap
33	62	73.8	329	15	US-10-225-067-28	Sequence 38, Appl
34	62	73.8	329	15	US-10-374-780A-214	Sequence 28, Appl
35	62	73.8	329	15	US-10-412-699B-700	Sequence 214, App
36	62	73.8	329	18	US-10-225-066A-1082	Sequence 700, App
37	62	73.8	476	15	US-10-264-049-2300	Sequence 1082, App
38	62	73.8	601	143	US-11-097-143-11721	Sequence 2300, Ap
39	62	73.8	831	9	US-09-789-561-86	Sequence 11721, A
40	62	73.8	831	11	US-09-833-245-2153	Sequence 86, Appl
41	62	73.8	831	17	US-10-883-936-86	Sequence 2153, Ap
42	62	73.8	835	16	US-10-723-860-4161	Sequence 86, Appl
43	62	73.8	835	18	US-10-756-149-5729	Sequence 4161, Ap
44	62	73.8	835	18	US-10-322-666-6	Sequence 5729, Ap
45	62	73.8	883	16	US-10-767-701-50588	Sequence 6, Appl1
46	61	72.6	178	16	US-10-424-599-224851	Sequence 50588, A
47	61	72.6	250	15	US-10-424-599-224851	Sequence 224851,
48	61	72.6	429	15	US-10-282-122A-52559	Sequence 52559, A
49	61	72.6	850	20	US-11-097-143-24084	Sequence 24084, A
50	61	72.6	1028	20	US-11-097-143-14916	Sequence 14916, A
51	59	70.2	38	15	US-10-424-599-246319	Sequence 246319,
52	59	70.2	45	15	US-10-424-599-150174	Sequence 150174,
53	59	70.2	59	18	US-10-450-763-36244	Sequence 36244, A
54	59	70.2	87	16	US-10-767-701-58439	Sequence 58439, A
55	59	70.2	124	18	US-10-450-763-43238	Sequence 43238, A
56	59	70.2	130	15	US-10-424-599-221693	Sequence 221693,
57	59	70.2	171	16	US-10-451-469A-414	Sequence 414, App
58	59	70.2	186	14	US-10-029-386-34005	Sequence 34005, A
59	59	70.2	217	15	US-10-424-599-153501	Sequence 153501,
60	59	70.2	281	15	US-10-282-122A-63467	Sequence 63467, A
61	59	70.2	292	9	US-09-864-761-37944	Sequence 37944, A
62	59	70.2	299	20	US-11-097-143-10602	Sequence 10602, A
63	59	70.2	374	20	US-11-097-143-8499	Sequence 8499, Ap
64	59	70.2	467	17	US-10-741-849-7304	Sequence 7304, Ap
65	59	70.2	537	20	US-11-097-143-4347	Sequence 4347, Ap
66	59	70.2	989	20	US-11-097-143-30167	Sequence 20661, A
67	59	70.2	1300	20	US-11-097-143-31017	Sequence 31017, A
68	58	69.0	51	18	US-10-450-763-40872	Sequence 40872, A
69	58	69.0	117	18	US-10-450-763-57203	Sequence 57203, A
70	58	69.0	156	9	US-09-864-761-41679	Sequence 41679, A
71	58	69.0	164	13	US-10-114-893-42	Sequence 42, Appl
72	58	69.0	324	18	US-10-450-763-50868	Sequence 50868, A
73	58	69.0	368	20	US-11-097-143-36393	Sequence 36393, A
74	58	69.0	378	14	US-10-029-386-33892	Sequence 33892, A
75	58	69.0	499	20	US-11-097-143-36015	Sequence 36015, A
76	58	69.0	695	20	US-11-097-143-2451	Sequence 2451, Ap
77	58	69.0	695	10	US-09-890-688-130	Sequence 130, App
78	58	69.0	749	20	US-11-097-143-23127	Sequence 23127, A
79	58	69.0	757	16	US-10-754-889-2	Sequence 2, Appl1
80	58	69.0	789	20	US-11-097-143-22004	Sequence 22004, A
81	57.5	68.5	178	15	US-10-424-599-175849	Sequence 175849,
82	57	67.9	188	15	US-10-424-599-193550	Sequence 193550,
83	57	67.9	193	15	US-10-424-599-169851	Sequence 169851,
84	57	67.9	227	18	US-10-450-763-44758	Sequence 44758, A

```
85 57 67.9 285 10 US-09-934-455-176 Sequence 176, App
86 57 67.9 285 14 US-10-278-173-16 Sequence 16, App
87 57 67.9 285 15 US-10-225-066A-56 Sequence 56, App
88 57 67.9 285 15 US-10-374-780A-226 Sequence 226, App
89 57 67.9 285 15 US-10-412-699B-726 Sequence 726, App
90 57 67.9 285 16 US-10-669-824-52 Sequence 52, App
91 57 67.9 285 17 US-10-870-198-52 Sequence 52, App
92 57 67.9 285 18 US-10-870-198-52 Sequence 52, App
93 57 67.9 485 20 US-11-097-143-5616 Sequence 5616, App
94 57 67.9 660 20 US-11-097-143-314-38 Sequence 38, App
95 57 67.9 660 20 US-11-097-143-314-38 Sequence 38, App
96 57 67.9 1146 17 US-10-741-849-7248 Sequence 7248, App
97 57 67.9 1161 20 US-11-097-143-38049 Sequence 38049, A
98 56 66.7 204 18 US-10-450-763-38565 Sequence 38565, A
99 56 66.7 212 16 US-10-451-467A-296 Sequence 296, App
100 56 66.7 407 14 US-10-087-887-76 Sequence 76, App
```

ALIGNMENTS

```
RESULT 1
US-09-858-332-18
; Sequence 18, Application US/09858332
; Patent No. US20020164718A1
; GENERAL INFORMATION:
; APPLICANT: Tchaga, Grigory S.
; TITLE OF INVENTION: Metal Ion Affinity Tags and Methods for
; FILE REFERENCE: CLON056CIP
; CURRENT APPLICATION NUMBER: US/09/858,332
; PRIOR FILING DATE: 2002-07-02, 017
; PRIOR APPLICATION NUMBER: 09/404,017
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,867
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-09-858-332-18
```

```
Query Match          100.0%; Score 84; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.7e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 HNNNNNNNNNN 12
DB      1 HNNNNNNNNNN 12
```

RESULT 2

```
US-10-453-410-4
; Sequence 4, Application US/10453410
; Publication No. US20040248101A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, YIN
; APPLICANT: TAN, XIN XING
; TITLE OF INVENTION: IDENTIFICATION OF NOVEL ANTIBACTERIAL AGENTS BY
; FILE REFERENCE: CRVA,025
; CURRENT APPLICATION NUMBER: US/10/453,410
; CURRENT FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; OTHER INFORMATION:
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-453-410-4
```

```
Query Match          100.0%; Score 84; DB 16; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.7e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 HNNNNNNNNNN 12
DB      1 HNNNNNNNNNN 12
```

RESULT 3

```
US-10-743-956-33
; Sequence 33, Application US/10743956
; Publication No. US20050136393A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, YIN
; APPLICANT: TAN, XIN XIN
; TITLE OF INVENTION: OLIGODEOXYNUCLEOTIDE (ODN) LIBRARIES, THEIR USE IN
; TITLE OF INVENTION: SCREENING FOR ANTIBACTERIAL AGENTS, AND CATALYTIC ODN
; FILE REFERENCE: CRVA,025/CIP
; CURRENT APPLICATION NUMBER: US/10/743,956
; PRIOR FILING DATE: 2003-12-23
; PRIOR APPLICATION NUMBER: 10/453,410
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 33
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic 6xN
US-10-743-956-33
```

```
Query Match          100.0%; Score 84; DB 18; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.7e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 HNNNNNNNNNN 12
DB      1 HNNNNNNNNNN 12
```

RESULT 4

```
US-10-055-794-15
; Sequence 15, Application US/10055794
; Publication No. US20030059900A1
; GENERAL INFORMATION:
; APPLICANT: Farmer, Alan Andrew
; TITLE OF INVENTION: SEQUENCE SPECIFIC RECOMBINASE-BASED
; TITLE OF INVENTION: METHODS FOR PRODUCING INTRON CONTAINING VECTORS AND
; FILE REFERENCE: CLON-069
; CURRENT APPLICATION NUMBER: US/10/055,794
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/263,358
; PRIOR FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: tag sequence
US-10-055-794-15
```



```

QY      1 HNNNNNNNNNN 12
      1111111111
Db      2 HNNNNNNNNNN 13

RESULT 5
US-10-390-472-53
; Sequence 53, Application US/10390472
; Publication No. US20040087013a1
; GENERAL INFORMATION:
APPLICANT: Holmes, Todd
          Zhang, Shuangang
          Rich, Alexander
          DiPersio, C. Michael
          Lockehin, Curtis
TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
                    SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
                    THEREFOR
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/390,472
FILING DATE: 17-Mar-2003
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/293,284
FILING DATE: 22-AUG-1994
APPLICATION NUMBER: 07/973,326
FILING DATE: 28-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 53:
US-10-390-472-53

Query Match      100.0%; Score 84; DB 15; Length 13;
Best Local Similarity 100.0%; Pred. No. 7,2e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HNNNNNNNNNN 12
      1111111111
Db      1 HNNNNNNNNNN 12

RESULT 6
US-10-422-262-16
; Sequence 16, Application US/10422262
Query Match      100.0%; Score 84; DB 15; Length 16;
Best Local Similarity 100.0%; Pred. No. 8,7e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HNNNNNNNNNN 12
      1111111111
Db      1 HNNNNNNNNNN 12

```

```

PUBLICatioN No. US20030219848A1
GENERAL INFORMATION:
APPLICANT: MAQVI, TABASSUM
APPLICANT: ROOHANI, RIJAZ
APPLICANT: SINGH, RAJENDRA
TITLE OF INVENTION: SHORT ENZYME DONOR FRAGMENT
FILE REFERENCE: 3817.11-1
CURRENT APPLICATION NUMBER: US/10/422,262
PRIOR FILING DATE: 2003-04-24
PRIOR APPLICATION NUMBER: 60/376,935
PRIOR FILING DATE: 2002-05-02
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 16
LENGTH: 55
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-422-262-16

Query Match          100.0%; Score 84; DB 15; Length 55;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
   |||||
Db 1 HNNNNNNNNNN 12

RESULT 7
US-10-448-609-12
; Sequence 12, Application US/10448609
; Publication No. US20040018562A1
GENERAL INFORMATION:
APPLICANT: MAQVI, TABASSUM
APPLICANT: ROOHANI, RIJAZ
APPLICANT: SINGH, RAJENDRA
TITLE OF INVENTION: IMPROVED RECEPTOR DETECTION
FILE REFERENCE: 3817.12-1
CURRENT APPLICATION NUMBER: US/10/448,609
PRIOR FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: 60/384,060
PRIOR FILING DATE: 2002-05-29
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 55
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-448-609-12

Query Match          100.0%; Score 84; DB 15; Length 55;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
   |||||
Db 1 HNNNNNNNNNN 12

RESULT 8
US-09-327-750-52
; Sequence 52, Application US/09327750
; Publication No. US20030079237A1
GENERAL INFORMATION:
APPLICANT: THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF
TITLE OF INVENTION: GENE ENCODING MADE, P75 NTR- ASSOCIATED CELL DEATH
TITLE OF INVENTION: EXECUTOR AND USES THEREOF

```

```
FILE REFERENCE: 59131apct
CURRENT APPLICATION NUMBER: US/09/327,750
CURRENT FILING DATE: 1999-06-07
NUMBER OF SEQ ID NOS: 56
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 52
LENGTH: 120
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-327-750-52
```

```
Query Match          94.0%; Score 79; DB 10; Length 120;
Best Local Similarity 91.7%; Pred. No. 0.0023;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 HNNHNNHNNH 12
        |||||
Db      29 HNNHNNHNNH 40
```

```
RESULT 9
US-09-826-752-4
Sequence 4, Application US/09826752
Patent No. US20010026930A1
GENERAL INFORMATION:
APPLICANT: Guarente, Leonard P.
APPLICANT: Austriaco Jr., Nicanor
APPLICANT: Claus, James J.
APPLICANT: Cole, Francesca
APPLICANT: Kennedy, Brian
TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESENCE IN
FILE REFERENCE: 0050.1491-005
CURRENT APPLICATION NUMBER: US/09/826,752
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: US 08/396,001
PRIOR FILING DATE: 1995-02-28
PRIOR APPLICATION NUMBER: PCT/US94/09351
PRIOR FILING DATE: 1994-08-15
PRIOR APPLICATION NUMBER: US 08/107,408
PRIOR FILING DATE: 1993-08-16
PRIOR APPLICATION NUMBER: US 09/323,433
PRIOR FILING DATE: 1999-06-01
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 834
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-826-752-4
```

```
Query Match          92.9%; Score 78; DB 9; Length 834;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 HNNHNNHNNH 11
        |||||
Db      609 HNNHNNHNNH 619
```

```
RESULT 10
US-10-912-434-4
Sequence 4, Application US/10912434
Publication No. US20050009169A1
GENERAL INFORMATION:
APPLICANT: Guarente, Leonard P.
APPLICANT: Austriaco Jr., Nicanor
APPLICANT: Claus, James J.
APPLICANT: Cole, Francesca
APPLICANT: Kennedy, Brian
TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESENCE IN
FILE REFERENCE: 0050.1491-005
```

```
CURRENT APPLICATION NUMBER: US/10/912,434
CURRENT FILING DATE: 2004-08-05
PRIOR APPLICATION NUMBER: US/09/826,752
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: US 08/396,001
PRIOR FILING DATE: 1995-02-28
PRIOR APPLICATION NUMBER: PCT/US94/09351
PRIOR FILING DATE: 1994-08-15
PRIOR APPLICATION NUMBER: US 08/107,408
PRIOR FILING DATE: 1993-08-16
PRIOR APPLICATION NUMBER: US 09/323,433
PRIOR FILING DATE: 1999-06-01
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 834
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-912-434-4
```

```
Query Match          92.9%; Score 78; DB 17; Length 834;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 HNNHNNHNNH 11
        |||||
Db      609 HNNHNNHNNH 619
```

```
RESULT 11
US-10-437-963-148051
Sequence 148051, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Bouharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 148051
LENGTH: 96
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MKT4530_48520C.1.pep
US-10-437-963-148051
```

```
Query Match          88.1%; Score 74; DB 16; Length 96;
Best Local Similarity 83.3%; Pred. No. 0.0083;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 HNNHNNHNNH 12
        |||||
Db      49 HHHHNNHNNH 60
```

```
RESULT 12
US-10-732-2998
Sequence 2998, Application US/10732923
Publication No. US20050108791A1
GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
FILE REFERENCE: 38-15(52796)C
```

CURRENT APPLICATION NUMBER: US/10/732.923
CURRENT FILING DATE: 2003-12-10
PRIOR APPLICATION NUMBER: 10/310.154
PRIOR FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 24149
SEQ ID NO 2398
LENGTH: 490
TYPE: PRT
ORGANISM: Dictyostelium discoideum
US-10-732-2998

Query Match 83.3%; Score 70; DB 17; Length 490;
Best Local Similarity 83.3%; Pred. No. 0.12;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:|:
DB 325 HNNNNNNNNNN 336

RESULT 13
US-10-424-599-200505
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Kovalic, David K
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424.599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 200505
LENGTH: 94
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_23081C.1.pep
US-10-424-599-200505

Query Match 82.1%; Score 69; DB 15; Length 94;
Best Local Similarity 75.0%; Pred. No. 0.036;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:|:
DB 24 HNNNNNNNNNN 35

RESULT 14
US-10-424-599-250721
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Kovalic, David K
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424.599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 250721
LENGTH: 105
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_68430C.1.pep

US-10-424-599-250721

Query Match 82.1%; Score 69; DB 15; Length 105;
Best Local Similarity 75.0%; Pred. No. 0.04;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:|:
DB 82 HNNNNNNNNNN 93

RESULT 15
US-10-437-963-152529
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boutharov, Andrey A.
APPLICANT: Barbazuk, Brad
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437.963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 152529
LENGTH: 193
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_52571C.1.pep
US-10-437-963-152529

Query Match 82.1%; Score 69; DB 16; Length 193;
Best Local Similarity 75.0%; Pred. No. 0.068;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:|:
DB 32 HSHSHSHNNNN 43

RESULT 16
US-10-296-144-5
Publication No. US20030113329A1
GENERAL INFORMATION:
APPLICANT: Bayer AG
TITLE OF INVENTION: REGULATION OF HUMAN TRANSCRIPTOLASE-LIKE ENZYME
FILE REFERENCE: L10078
CURRENT APPLICATION NUMBER: US/10/296.144
CURRENT FILING DATE: 2002-11-22
PRIOR FILING DATE: 2000-05-31
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 1284
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-10-296-144-5

Query Match 82.1%; Score 69; DB 14; Length 1284;
Best Local Similarity 75.0%; Pred. No. 0.37;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
|:|:|:|:|:|:|:

Db 1163 HHHNHHSHNHN 1174

RESULT 17
US-11-097-143-42399

Sequence 42399, Application US/11097143
Publication No. US20050208558A1

GENERAL INFORMATION:

APPLICANT: Venter, J. Craig

APPLICANT: et al.

TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID

TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE

FILE REFERENCE: DROSOPHILA GENES.

FILE REFERENCE: CLO00728

CURRENT APPLICATION NUMBER: US/11/097,143

CURRENT FILING DATE: 2005-04-04

PRIOR APPLICATION NUMBER: 60/157,832

PRIOR FILING DATE: 1999-10-05

PRIOR APPLICATION NUMBER: 60/160,191

PRIOR FILING DATE: 1999-10-19

PRIOR APPLICATION NUMBER: 60/161,932

PRIOR FILING DATE: 1999-10-28

PRIOR APPLICATION NUMBER: 60/164,769

PRIOR FILING DATE: 1999-11-12

PRIOR APPLICATION NUMBER: 60/173,383

PRIOR FILING DATE: 1999-12-28

PRIOR APPLICATION NUMBER: 60/175,693

PRIOR FILING DATE: 2000-01-12

PRIOR APPLICATION NUMBER: 60/184,831

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: 60/191,637

PRIOR FILING DATE: 2000-03-23

NUMBER OF SEQ ID NOS: 43008

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 42399

LENGTH: 1284

TYPE: PRT

ORGANISM: DROSOPHILA

US-11-097-143-42399

Query Match Best Local Similarity 82.1%; Score 69; DB 20; Length 1284;

Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHHNHHNHNHN 12

Db 1163 HHHNHHSHNHN 1174

RESULT 18
US-11-097-143-21195

Sequence 21195, Application US/11097143

Publication No. US20050208558A1

GENERAL INFORMATION:

APPLICANT: Venter, J. Craig

APPLICANT: et al.

TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID

TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE

FILE REFERENCE: CLO00728

CURRENT APPLICATION NUMBER: US/11/097,143

CURRENT FILING DATE: 2005-04-04

PRIOR APPLICATION NUMBER: 60/157,832

PRIOR FILING DATE: 1999-10-05

PRIOR APPLICATION NUMBER: 60/160,191

PRIOR FILING DATE: 1999-10-19

PRIOR APPLICATION NUMBER: 60/161,932

PRIOR FILING DATE: 1999-10-28

PRIOR APPLICATION NUMBER: 60/164,769

PRIOR FILING DATE: 1999-11-12

PRIOR APPLICATION NUMBER: 60/173,383

PRIOR FILING DATE: 1999-12-28

PRIOR APPLICATION NUMBER: 60/175,693

PRIOR FILING DATE: 2000-01-12

PRIOR APPLICATION NUMBER: 60/184,831

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: 60/191,637

PRIOR FILING DATE: 2000-03-23

NUMBER OF SEQ ID NOS: 43008

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 21195

LENGTH: 697

TYPE: PRT

ORGANISM: DROSOPHILA

US-11-097-143-21195

Query Match Best Local Similarity 81.5%; Score 68.5; DB 20; Length 697;

Matches 11; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 HHHNH-NHNHNHN 12

Db 622 HHHNHHNHNHNHS 634

RESULT 19
US-10-390-472-60

Sequence 60, Application US/10390472

Publication No. US20040087013A1

GENERAL INFORMATION:

APPLICANT: Holmes, Todd

Zhang, Shuguang

Rich, Alexander

Dipersio, C. Michael

Locke, Curtiss

TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY

SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES

THEORETICAL

NUMBER OF SEQUENCES: 64

CORRESPONDENCE ADDRESS:

ADDRESS: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02173-4799

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/390,472

FILING DATE: 17-Mar-2003

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/293,284

FILING DATE: 22-AUG-1994

APPLICATION NUMBER: 07/973,326

FILING DATE: 28-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Brook, David E.

REGISTRATION NUMBER: 22,592

REFERENCE/DOCKET NUMBER: MIT-6008A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 861-6240

TELEFAX: (617) 861-9540

INFORMATION FOR SEQ ID NO: 60:

SEQUENCE CHARACTERISTICS:

LENGTH: 16 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 60:

US-10-390-472-60

Query Match 78.6%; Score 66; DB 15; Length 16;
Best Local Similarity 81.8%; Pred. No. 0.018;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
| | | | | | | |
Db 1 HNNHNNHNNH 11

RESULT 20

US-10-390-472-61
; Sequence 61, Application US/10390472
; Publication No. US20040087013A1
; GENERAL INFORMATION:

APPLICANT: Holmes, Todd

Rich, Shunguang
Dipersio, C. Michael
Lockshin, Curtis

TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
THEREFOR

NUMBER OF SEQUENCES: 64

CORRESPONDENCE ADDRESSES:

ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.

STREET: Two Millitia Drive

CITY: Lexington

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02173-4799

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION NUMBER: US/10/390,472

APPLICATION NUMBER: 435

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/293,284

FILING DATE: 22-AUG-1994

APPLICATION NUMBER: 07/973,326

FILING DATE: 28-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Brook, David E.

REGISTRATION NUMBER: 22,592

REFERENCE/DOCKET NUMBER: MIT-6008A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 861-6240

TELEFAX: (617) 861-9540

INFORMATION FOR SEQ ID NO: 61:

SEQUENCE CHARACTERISTICS:

LENGTH: 16 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 61:

Query Match 78.6%; Score 66; DB 15; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.018;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
| | | | | | | |
Db 1 HNNHNNHNNH 12

RESULT 21

US-10-425-115-204424
; Sequence 204424, Application US/10425115
; Publication No. US20040214272A1

GENERAL INFORMATION:
; APPLICANT: la Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

PLANTS
FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 204424

LENGTH: 148

TYPE: PRT

ORGANISM: Zea mays

FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_118026C.1.pep

US-10-425-115-204424

Query Match 78.6%; Score 66; DB 16; Length 148;
Best Local Similarity 81.8%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
| | | | | | | |
Db 1 HNNHNNHNNH 11

RESULT 22

US-09-327-750-34
; Sequence 34, Application US/09327750
; Publication No. US20030079237A1
; GENERAL INFORMATION:

APPLICANT: THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF

TITLE OF INVENTION: GENE ENCODING MADE, P75 NTR- ASSOCIATED CELL DEATH

FILE REFERENCE: 59131apct

CURRENT APPLICATION NUMBER: US/09/327,750

CURRENT FILING DATE: 1999-06-07

NUMBER OF SEQ ID NOS: 56

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 34

LENGTH: 124

TYPE: PRT

ORGANISM: Mus musculus

US-09-327-750-34

Query Match 77.4%; Score 65; DB 10; Length 124;
Best Local Similarity 75.0%; Pred. No. 0.15;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
| | | | | | | |
Db 37 NNNHNNHNNH 48

RESULT 23

US-09-327-750-53
; Sequence 53, Application US/09327750
; Publication No. US20030079237A1
; GENERAL INFORMATION:

APPLICANT: THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF

TITLE OF INVENTION: GENE ENCODING MADE, P75 NTR- ASSOCIATED CELL DEATH

FILE REFERENCE: 59131apct

CURRENT APPLICATION NUMBER: US/09/327,750

CURRENT FILING DATE: 1999-06-07

NUMBER OF SEQ ID NOS: 56

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 53

LENGTH: 124

TYPE: PRT

ORGANISM: Mus musculus

US-09-327-750-53

Query Match 77.4%; Score 65; DB 10; Length 124;
Best Local Similarity 75.0%; Pred. No. 0.15;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
:|:|:|:|:|:|:
Db 37 NNNNNHNNH 48

RESULT 24

US-10-424-599-180331
; Sequence 180331, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 180331
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_133854C.1.pap
US-10-424-599-180331

Query Match 76.2%; Score 64; DB 15; Length 276;
Best Local Similarity 66.7%; Pred. No. 0.41;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
|:|:|:|:|:|:
Db 48 HNNHNNHNNH 59

RESULT 25

US-10-282-122A-47210
; Sequence 47210, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cair, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47210
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Botryella burgdorferi
US-10-282-122A-47210

Query Match 76.2%; Score 64; DB 15; Length 451;
Best Local Similarity 66.7%; Pred. No. 0.64;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
|:|:|:|:|:|:
Db 253 HDHDDHNNH 264

RESULT 26

US-10-739-930-5773
; Sequence 5773, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 5773
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-C1379_1.p
US-10-739-930-5773

Query Match 76.2%; Score 64; DB 16; Length 545;
Best Local Similarity 75.0%; Pred. No. 0.76;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
:|:|:|:|:|:|:
Db 474 NNNHNNHNNH 485

RESULT 27

US-10-041-018-267
; Sequence 267, Application US/10041018
; Publication No. US20040072323A1
; GENERAL INFORMATION:
; APPLICANT: Matsuda, Seiji P.T.
; APPLICANT: Hart, Elizabeth A.
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism
; FILE REFERENCE: P02080US1/10025547
; CURRENT APPLICATION NUMBER: US/10/041,018
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/259880
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: PatentIn version 3.1

SEQ ID NO 267
LENGTH: 1176
TYPE: PRT
ORGANISM: Phycomyces blakesleeanus
US-10-041-018-267

Query Match 76.2%; Score 64; DB 15; Length 1176;
Best Local Similarity 66.7%; Pred. No. 1.5;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 12
|:|:|:|:|:|:
Db 705 HNNHSHSHSHN 716

RESULT 28
US-09-978-360A-708
Sequence 708, Application US/09978360A
GENERAL INFORMATION:
APPLICANT: Duclet, Jean-Baptiste Dumas Milne
APPLICANT: Duclet, Aymeric
APPLICANT: Bouquellet, Lydie
APPLICANT: Jober, Severin
APPLICANT: Clusel, Catherine
TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
FILE REFERENCE: 56, US4, CIP
CURRENT APPLICATION NUMBER: US/09/978,360A
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: US 60/066,677
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: US 60/069,957
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: US 60/074,121
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: US 60/081,563
PRIOR FILING DATE: 1998-04-13
PRIOR APPLICATION NUMBER: US 60/096,116
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: US 60/099,273
PRIOR FILING DATE: -09-04
PRIOR APPLICATION NUMBER: US 09/191,997
PRIOR FILING DATE: 1998-11-13
PRIOR APPLICATION NUMBER: US 09/215,435
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: PCT/IB98/02122
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: US 09/247,155
PRIOR FILING DATE: 1999-02-09
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 810
SOFTWARE: Patent.pm
SEQ ID NO 708
LENGTH: 150
TYPE: PRT
ORGANISM: Homo sapiens
US-09-978-360A-708

Query Match 73.8%; Score 62; DB 11; Length 150;
Best Local Similarity 72.7%; Pred. No. 0.43;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 11
|:|:|:|:|:|:
Db 124 HSHNHSHSHN 134

RESULT 29
US-10-319-763-196
Sequence 196, Application US/10319763
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean-Baptiste

APPLICANT: Duclet, Aymeric
APPLICANT: Bouquellet, Lydie
TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
FILE REFERENCE: G-031, US04, DIV
CURRENT APPLICATION NUMBER: US/10/319,763
CURRENT FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: 60/066,677
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/069,957
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/074,121
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/081,563
PRIOR FILING DATE: 1998-04-13
PRIOR APPLICATION NUMBER: 60/096,116
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/099,273
PRIOR FILING DATE: 1998-09-04
NUMBER OF SEQ ID NOS: 229
SOFTWARE: Patent.pm
SEQ ID NO 196
LENGTH: 150
TYPE: PRT
ORGANISM: Homo sapiens
US-10-319-763-196

Query Match 73.8%; Score 62; DB 14; Length 150;
Best Local Similarity 72.7%; Pred. No. 0.43;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HNNHNNHNNH 11
|:|:|:|:|:|:
Db 124 HSHNHSHSHN 134

RESULT 30
US-09-934-455-166
Sequence 166, Application US/09934455
Publication No. US20030121070A1
GENERAL INFORMATION:
APPLICANT: Adam, Luc
APPLICANT: Creelman, Robert
APPLICANT: Dubell, Arnold
APPLICANT: Heard, Jacqueline
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Kedde, James
APPLICANT: Pilgrim, Marsha
APPLICANT: Ratcliffe, Oliver
APPLICANT: Reuber, Lynne
APPLICANT: Riechmann, Jose Luis
APPLICANT: Yu, Guo-Liang
APPLICANT: Plineda, Omaira
TITLE OF INVENTION: Genes for Modifying Plant Traits IV
FILE REFERENCE: MBI-0025
CURRENT APPLICATION NUMBER: US/09/934,455
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/227439
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: MBI-0022
PRIOR FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: MBI-0023
PRIOR FILING DATE: 2001-04-17
NUMBER OF SEQ ID NOS: 516
SOFTWARE: PatentIn version 3.1
SEQ ID NO 166
LENGTH: 329
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-934-455-166

Query Match 73.8%; Score 62; DB 10; Length 329;
Best Local Similarity 66.7%; Pred. No. 0.87;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 HNNNNNNNNNN 12
      ||:|||||:|:|:|
Db      101 HNNNNHHHHHS 112

RESULT 31
US-10-286-264-24
/ Sequence 24, Application US/10286264
/ Publication No. US2003093837A1
/ GENERAL INFORMATION:
/ APPLICANT: Keddie, James
/ APPLICANT: Riechmann, Jose-Luis
/ APPLICANT: Ratcliffe, Oliver
/ APPLICANT: Zhang, James
/ APPLICANT: Jiang, Cai-Zhong
/ APPLICANT: Pineda, Omaisra
/ APPLICANT: Heard, Jacqueline
/ APPLICANT: Yu, Guo-Liang
/ APPLICANT: Adam, Luc
/ APPLICANT: Broun, Pierre
/ APPLICANT: Reuber, Lynne
/ APPLICANT: Pilgrim, Marsha
/ APPLICANT: Samaha, Raymond
/ TITLE OF INVENTION: POLYNUCLEOTIDES FOR SEED TRAIT ALTERATION
/ FILE REFERENCE: MBI-008
/ CURRENT APPLICATION NUMBER: US/10/286,264
/ CURRENT FILING DATE: 2002-11-01
/ PRIOR APPLICATION NUMBER: 60/125,814
/ PRIOR FILING DATE: 1999-03-23
/ NUMBER OF SEQ ID NOS: 165
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 24
/ LENGTH: 329
/ TYPE: PRT
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ OTHER INFORMATION: G1052
US-10-286-264-24

Query Match      73.8%; Score 62; DB 14; Length 329;
Best Local Similarity 66.7%; Pred. No. 0.87;
Matches      8; Conservative      4; Mismatches      0; Indels      0; Gaps      0;

QY      1 HNNNNNNNNNN 12
      ||:|||||:|:|:|
Db      101 HNNNNHHHHHS 112

RESULT 32
US-10-225-066A-1082
/ Sequence 1082, Application US/10225066A
/ Publication No. US20030226173A1
/ GENERAL INFORMATION:
/ APPLICANT: Mendel Biotechnology, Inc.
/ APPLICANT: Ratcliffe, Oliver
/ APPLICANT: Riechmann, Jose Luis
/ APPLICANT: Adam, Luc J
/ APPLICANT: DUBELL, Arnold T
/ APPLICANT: HEARD, Jacqueline E
/ APPLICANT: PILGRIM, Marsha L
/ APPLICANT: JIANG, Cai-Zhong
/ APPLICANT: REUBER, T. Lynne
/ APPLICANT: CREELMAN, Robert A
/ APPLICANT: PINEDA, Omaisra
/ APPLICANT: YU, Guo-Liang
/ APPLICANT: BROUN, Pierre E
/ TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
/ FILE REFERENCE: MBI0036-2 US
/ CURRENT APPLICATION NUMBER: US/10/225,066A
/ CURRENT FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 09/837,444
/ PRIOR FILING DATE: 2001-04-18
```

```
      / PRIOR APPLICATION NUMBER: 60/310,847
      / PRIOR FILING DATE: 2001-08-09
      / PRIOR APPLICATION NUMBER: 60/336,049
      / PRIOR FILING DATE: 2001-12-05
      / PRIOR APPLICATION NUMBER: 60/338,692
      / PRIOR FILING DATE: 2001-12-11
      / PRIOR APPLICATION NUMBER: 10/171,468
      / PRIOR FILING DATE: 2002-06-14
      / NUMBER OF SEQ ID NOS: 1122
      / SOFTWARE: PatentIn version 3.1
      / SEQ ID NO 1082
      / LENGTH: 329
      / TYPE: PRT
      / ORGANISM: Arabidopsis thaliana
US-10-225-066A-1082

Query Match      73.8%; Score 62; DB 15; Length 329;
Best Local Similarity 66.7%; Pred. No. 0.87;
Matches      8; Conservative      4; Mismatches      0; Indels      0; Gaps      0;

QY      1 HNNNNNNNNNN 12
      ||:|||||:|:|:|
Db      101 HNNNNHHHHHS 112

RESULT 33
US-10-302-267-38
/ Sequence 38, Application US/10302267
/ Publication No. US20030229915A1
/ GENERAL INFORMATION:
/ APPLICANT: Keddie, James
/ APPLICANT: Fromme, Michael
/ APPLICANT: Heard, Jacqueline
/ APPLICANT: Riechmann, Jose Luis
/ APPLICANT: Adam, Luc
/ APPLICANT: Broun, Pierre
/ APPLICANT: Pineda, Omaisra
/ APPLICANT: Reuber, Lynne
/ APPLICANT: Zhang, James
/ APPLICANT: Yu, Guo-Liang
/ APPLICANT: Jiang, Cai-Zhong
/ APPLICANT: Samaha, Raymond
/ APPLICANT: Pilgrim, Marsha
/ APPLICANT: Creelman, Robert
/ TITLE OF INVENTION: PLANT GENE SEQUENCES II
/ FILE REFERENCE: MBI-0007
/ CURRENT APPLICATION NUMBER: US/10/302,267
/ CURRENT FILING DATE: 2002-11-22
/ PRIOR APPLICATION NUMBER: US/09/506,720
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/120,880
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: 60/121,037
/ PRIOR FILING DATE: 1999-02-22
/ PRIOR APPLICATION NUMBER: 60/124,278
/ PRIOR FILING DATE: 1999-03-11
/ PRIOR APPLICATION NUMBER: 60/129,450
/ PRIOR FILING DATE: 1999-04-15
/ PRIOR APPLICATION NUMBER: 60/135,134
/ PRIOR FILING DATE: 1999-05-20
/ PRIOR APPLICATION NUMBER: 60/144,153
/ PRIOR FILING DATE: 1999-07-15
/ PRIOR APPLICATION NUMBER: 60/161,143
/ PRIOR FILING DATE: 1999-10-22
/ PRIOR APPLICATION NUMBER: 60/162,656
/ PRIOR FILING DATE: 1999-11-01
/ NUMBER OF SEQ ID NOS: 218
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 38
/ LENGTH: 329
/ TYPE: PRT
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
```



```
/ OTHER INFORMATION: G1052
US-10-302-267-38

Query Match      73.8%; Score 62; DB 15; Length 329;
Best Local Similarity 66.7%; Pred. No. 0.87;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 HNNNNNNNNNN 12
       ||:||||:|:|:|
Db      101 HNNNNHHHHHS 112

RESULT 34
US-10-225-067-28
; Sequence 28, Application US/10225067
; Publication No. US20040019925A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Creelman, Robert A.
; APPLICANT: Keddie, James
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Dubell, Arnold T.
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pineda, Omaisra
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Broun, Pierre E.
; TITLE OF INVENTION: BIOCHEMISTRY-RELATED POLYNUCLEOTIDES AND
; FILE REFERENCE: 514442002042
; CURRENT APPLICATION NUMBER: US/10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (201)...(261)
; OTHER INFORMATION: Conserved domain
US-10-225-067-28

Query Match      73.8%; Score 62; DB 15; Length 329;
Best Local Similarity 66.7%; Pred. No. 0.87;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 HNNNNNNNNNN 12
       ||:||||:|:|:|
Db      101 HNNNNHHHHHS 112

RESULT 35
US-10-374-780A-214
; Sequence 214, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Haake, Volker
```

```
/ APPLICANT: Creelman, Robert A
/ APPLICANT: Ratcliffe, Oliver
/ APPLICANT: Adam, Luc J.
/ APPLICANT: Reuber, T. Lynne
/ APPLICANT: Keddie, James
/ APPLICANT: Broun, Pierre E.
/ APPLICANT: Pilgrim, Marsha L.
/ APPLICANT: Dubell III, Arnold T.
/ APPLICANT: Pineda, Omaisra
/ APPLICANT: Yu, Guo-Liang
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
/ FILE REFERENCE: MB1-0047 CIP
/ CURRENT APPLICATION NUMBER: US/10/374,780A
/ PRIOR FILING DATE: 2003-02-25
/ PRIOR APPLICATION NUMBER: 09/837,944
/ PRIOR FILING DATE: 2001-04-18
/ PRIOR APPLICATION NUMBER: 60/310,847
/ PRIOR FILING DATE: 2001-08-09
/ PRIOR APPLICATION NUMBER: 09/934,455
/ PRIOR FILING DATE: 2001-08-22
/ PRIOR APPLICATION NUMBER: 60/336,049
/ PRIOR FILING DATE: 2001-11-19
/ PRIOR APPLICATION NUMBER: 60/338,692
/ PRIOR FILING DATE: 2001-12-11
/ PRIOR APPLICATION NUMBER: 10/171,468
/ PRIOR FILING DATE: 2002-06-14
/ PRIOR APPLICATION NUMBER: 10/225,066
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,067
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,068
/ PRIOR FILING DATE: 2002-08-09
/ NUMBER OF SEQ ID NOS: 2906
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 214
/ LENGTH: 329
/ TYPE: PRT
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ OTHER INFORMATION: G1052 (conserved domain in AA coordinates: 201-261)
US-10-374-780A-214

Query Match      73.8%; Score 62; DB 15; Length 329;
Best Local Similarity 66.7%; Pred. No. 0.87;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 HNNNNNNNNNN 12
       ||:||||:|:|:|
Db      101 HNNNNHHHHHS 112

RESULT 36
US-10-412-699B-700
; Sequence 700, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omaisra
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: Dubell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
```

```

; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 700
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1052
US-10-412-699B-700

```

```

Query Match          73.8%; Score 62; DB 15; Length 329;
Best Local Similarity 66.7%; Pred. No. 0.87;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 HNNNNNNNNNN 12
        ||:||||:|:|:|
Db      101 HNNNNHHHHS 112

```

```

RESULT 37
US-10-225-066A-1082
; Sequence 1082, Application US/10225066A
; Publication No. US20050160493A9
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DOBELT, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lyne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROUN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI0036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049

```

```

; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1082
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-1082

```

```

Query Match          73.8%; Score 62; DB 18; Length 329;
Best Local Similarity 66.7%; Pred. No. 0.87;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 HNNNNNNNNNN 12
        ||:||||:|:|:|
Db      101 HNNNNHHHHS 112

```

```

RESULT 38
US-10-264-049-2300
; Sequence 2300, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2300
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (377)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (437)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (470)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (474)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-2300

```

```

Query Match          73.8%; Score 62; DB 15; Length 476;
Best Local Similarity 72.7%; Pred. No. 1.2;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 HNNNNNNNNNN 11
        ||:||||:|:|:|
Db      133 HSHNHQSHNH 143

```

```

RESULT 39
US-11-097-143-11721
; Sequence 11721, Application US/11097143
; Publication No. US20050208558A1

```

```
/ GENERAL INFORMATION:
/ APPLICANT: Venter, J. Craig
/ APPLICANT: et al.
/ TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
/ TITLE OF INVENTION: ARAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
/ TITLE OF INVENTION: DROSOPHILA GENES.
/ FILE REFERENCE: CL000728
/ CURRENT APPLICATION NUMBER: US/11/097,143
/ CURRENT FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: 60/157,832
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/160,191
/ PRIOR FILING DATE: 1999-10-19
/ PRIOR APPLICATION NUMBER: 60/161,932
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: 60/164,769
/ PRIOR FILING DATE: 1999-11-12
/ PRIOR APPLICATION NUMBER: 60/173,383
/ PRIOR FILING DATE: 1999-12-28
/ PRIOR APPLICATION NUMBER: 60/175,693
/ PRIOR FILING DATE: 2000-01-12
/ PRIOR APPLICATION NUMBER: 60/184,831
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/191,637
/ PRIOR FILING DATE: 2000-03-23
/ NUMBER OF SEQ ID NOS: 4308
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 11721
/ LENGTH: 601
/ TYPE: PRT
/ ORGANISM: DROSOPHILA
/ US-11-097-143-11721
```

```
Query Match          73.8%; Score 62; DB 20; Length 601;
Best Local Similarity 66.7%; Pred. No. 1.5;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 HNNNNNNNNH 12
         |||||
Db      250 HHHHHHHNNN 261
```

```
RESULT 40
US-09-789-561-86
/ Sequence 86, Application US/09789561
/ Patent No. US20020064818A1
/ GENERAL INFORMATION:
/ APPLICANT: Ni et al.
/ TITLE OF INVENTION: 52 Human secreted proteins
/ FILE REFERENCE: P2043P1
/ CURRENT APPLICATION NUMBER: US/09/789,561
/ CURRENT FILING DATE: 2001-02-22
/ PRIOR APPLICATION NUMBER: PCT/US00/24008
/ PRIOR FILING DATE: 2000-08-31
/ PRIOR APPLICATION NUMBER: 60/152,317
/ PRIOR FILING DATE: 1999-09-03
/ PRIOR APPLICATION NUMBER: 60/152,315
/ PRIOR FILING DATE: 1999-09-03
/ NUMBER OF SEQ ID NOS: 194
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 86
/ LENGTH: 831
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-789-561-86
```

```
Query Match          73.8%; Score 62; DB 9; Length 831;
Best Local Similarity 72.7%; Pred. No. 2;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 HNNNNNNNNH 11
         |||||
Db      124 HSHNHQSHNH 134
```

```
RESULT 41
US-09-833-245-2153
/ Sequence 2153, Application US/09833245
/ Publication No. US20040010134A1
/ GENERAL INFORMATION:
/ APPLICANT: Human Genome Sciences, Inc.
/ TITLE OF INVENTION: Albumin Fusion Proteins
/ FILE REFERENCE: PP546PCT
/ CURRENT APPLICATION NUMBER: US/09/833,245
/ CURRENT FILING DATE: 2001-04-12
/ PRIOR APPLICATION NUMBER: 60/229,358
/ PRIOR FILING DATE: 2000-04-12
/ PRIOR APPLICATION NUMBER: 60/256,931
/ PRIOR FILING DATE: 2000-12-21
/ PRIOR APPLICATION NUMBER: 60/199,384
/ PRIOR FILING DATE: 2000-04-25
/ NUMBER OF SEQ ID NOS: 2267
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 2153
/ LENGTH: 831
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-833-245-2153
```

```
Query Match          73.8%; Score 62; DB 11; Length 831;
Best Local Similarity 72.7%; Pred. No. 2;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 HNNNNNNNNH 11
         |||||
Db      124 HSHNHQSHNH 134
```

```
RESULT 42
US-10-883-936-86
/ Sequence 86, Application US/10883936
/ Publication No. US20050019866A1
/ GENERAL INFORMATION:
/ APPLICANT: Ni et al.
/ TITLE OF INVENTION: 52 Human secreted proteins
/ FILE REFERENCE: P2043P1
/ CURRENT APPLICATION NUMBER: US/10/883,936
/ CURRENT FILING DATE: 2004-07-06
/ PRIOR APPLICATION NUMBER: US/09/789,561
/ PRIOR FILING DATE: 2001-02-22
/ PRIOR APPLICATION NUMBER: PCT/US00/24008
/ PRIOR FILING DATE: 2000-08-31
/ PRIOR APPLICATION NUMBER: 60/152,317
/ PRIOR FILING DATE: 1999-09-03
/ PRIOR APPLICATION NUMBER: 60/152,315
/ PRIOR FILING DATE: 1999-09-03
/ NUMBER OF SEQ ID NOS: 194
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 86
/ LENGTH: 831
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-10-883-936-86
```

```
Query Match          73.8%; Score 62; DB 17; Length 831;
Best Local Similarity 72.7%; Pred. No. 2;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 HNNNNNNNNH 11
         |||||
Db      124 HSHNHQSHNH 134
```

```
RESULT 43
US-10-723-860-4161
/ Sequence 4161, Application US/10723860
```

Publication No. US20040253606A1
GENERAL INFORMATION:
APPLICANT: Aziz, Natasha
APPLICANT: Ginsburg, Wendy M.
TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
FILE REFERENCE: 05882.0193.NUS01
CURRENT APPLICATION NUMBER: US/10/723, 860
CURRENT FILING DATE: 2003-11-26
PRIORITY APPLICATION NUMBER: 60/429,739
PRIORITY FILING DATE: 2002-11-26
NUMBER OF SEQ ID NOS: 8393
SOFTWARE: PatentIn version 3.2
SEQ ID NO 4161
LENGTH: 835
TYPE: PRT
ORGANISM: Homo sapiens
US-10-723-860-4161

Query Match 73.8%; Score 62; DB 16; Length 835;
Best Local Similarity 72.7%; Pred. No. 2;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNNNNNNNH 11
|:|:|:|:|:|:|
Db 128 HSHNHQSHNH 138

RESULT 44
US-10-756-149-5729
Sequence 5729, Application US/10756149
Publication No. US20050181375A1
GENERAL INFORMATION:
APPLICANT: Aziz, Natasha
APPLICANT: Zlotnick, Albert
TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
FILE REFERENCE: file
CURRENT APPLICATION NUMBER: US/10/756,149
CURRENT FILING DATE: 2004-01-12
NUMBER OF SEQ ID NOS: 5818
SOFTWARE: PatentIn version 3.2
SEQ ID NO 5729
LENGTH: 835
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-756-149-5729

Query Match 73.8%; Score 62; DB 18; Length 835;
Best Local Similarity 72.7%; Pred. No. 2;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNNNNNNNH 11
|:|:|:|:|:|:|
Db 128 HSHNHQSHNH 138

RESULT 45
US-10-322-696-6
Sequence 6, Application US/10322696
Publication No. US20040166490A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Malandro, Marc
TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
FILE REFERENCE: 529452001200
CURRENT APPLICATION NUMBER: US/10/322,696
CURRENT FILING DATE: 2003-10-17
NUMBER OF SEQ ID NOS: 186
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 883

TYPE: PRT
ORGANISM: Homo sapiens
US-10-322-696-6

Query Match 73.8%; Score 62; DB 16; Length 883;
Best Local Similarity 72.7%; Pred. No. 2.1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNNNNNNNH 11
|:|:|:|:|:|:|
Db 176 HSHNHQSHNH 186

RESULT 46
US-10-767-701-50588
Sequence 50588, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovacic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(5353)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 50588
LENGTH: 178
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(178)
OTHER INFORMATION: unsure at all Xaa locations
OTHER INFORMATION: Clone ID: LIB3480-018-P1-K1-A10.pcp
US-10-767-701-50588

Query Match 72.6%; Score 61; DB 16; Length 178;
Best Local Similarity 81.8%; Pred. No. 0.66;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNNNNNNNNH 12
|:|:|:|:|:|:|
Db 62 NNNNNHYNNH 72

RESULT 47
US-10-424-599-224861
Sequence 224861, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Zhou Yihua
APPLICANT: Kovacic David K
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 224861
LENGTH: 250
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_45080C.1.pcp
US-10-424-599-224861

Query Match 72.6%; Score 61; DB 15; Length 250;
Best Local Similarity 90.0%; Pred. No. 0.92;

```
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 NNNNNNNNN 11
    |||||
Db 12 NNNNNNNNN 21

RESULT 48
US-10-282-122A-52569
; Sequence 52569, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyckind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52569
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-10-282-122A-52569

Query Match 72.6%; Score 61; DB 15; Length 429;
Best Local Similarity 66.7%; Pred. No. 1.5;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24084
; LENGTH: 850
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-24084

Query Match 72.6%; Score 61; DB 20; Length 850;
Best Local Similarity 66.7%; Pred. No. 2.7;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 NNNNNNNNN 12
    |||||
Db 144 NNNNNNNNN 155

RESULT 50
US-11-097-143-14916
; Sequence 14916, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14916
; LENGTH: 1028
; TYPE: PRT
```

ORGANISM: DROSOPHILA
US-11-097-143-14916

Query Match
Best Local Similarity 72.6%; Score 61; DB 20; Length 1028;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|||:|:|:|:
DB 144 HNNHHHHQHNNN 155

RESULT 51

US-10-424-599-246199
; Sequence 246199, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 246199
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_6434C.1.dep
US-10-424-599-246199

Query Match
Best Local Similarity 70.2%; Score 59; DB 15; Length 38;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|||:|:|:|:
DB 9 HNNHHHHHNNH 20

RESULT 52

US-10-424-599-150174
; Sequence 150174, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 150174
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(45)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_10662C.1.dep
US-10-424-599-150174

Query Match
Best Local Similarity 70.2%; Score 59; DB 15; Length 45;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
|||:|:|:|:
DB 33 HNNHHHHLNNH 43

RESULT 53

US-10-450-763-36244
; Sequence 36244, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyeq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 36244
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (23)..(47)
; OTHER INFORMATION: HMW KININOGEN SIGNATURE domain identified by eMATRIX,
US-10-450-763-36244

Query Match
Best Local Similarity 70.2%; Score 59; DB 18; Length 59;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|||:|:|:|:
DB 20 HNNHHHHHNNH 31

RESULT 54

US-10-767-701-58439
; Sequence 58439, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535) B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 58439
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30979644.dep
US-10-767-701-58439

Query Match
Best Local Similarity 70.2%; Score 59; DB 16; Length 87;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|||:|:|:|:

Db 68 HHHHHHHHHH 79

```
RESULT 55
US-10-450-763-43238
; Sequence 43238, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450.763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 43238
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (68)..(92)
; OTHER INFORMATION: HMW KININOGEN SIGNATURE domain identified by eMATRIX,
US-10-450-763-43238
; accession number PR0034B, p-value=2.800e-10, raw score of 8.69
```

Query Match 70.2%; Score 59; DB 18; Length 124;
Best Local Similarity 58.3%; Pred. No. 0.88;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHHHHHHHHH 12
|:|:|:|:|:|:|

```
RESULT 56
US-10-424-599-221693
; Sequence 221693, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 221693
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_42218C.1.pep
US-10-424-599-221693
```

Query Match 70.2%; Score 59; DB 15; Length 130;
Best Local Similarity 66.7%; Pred. No. 0.92;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HHHHHHHHHH 12
|:|:|:|:|:|:|

Db 22 HHHHHHHHHH 33

RESULT 57

```
US-10-451-467A-414
; Sequence 414, Application US/10451467A
; Publication No. US20040161840A1
; GENERAL INFORMATION:
; APPLICANT: CONTRERAS, ROLAND HENRI
; APPLICANT: BEERHARDT, INES
; APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS
; APPLICANT: REEKMAN, RIEKA JOSEPHINA
; TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN
; TITLE OF INVENTION: YEAST AND FUNGI
; FILE REFERENCE: JAB-1667
; CURRENT APPLICATION NUMBER: US/10/451,467A
; CURRENT FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: EP 00870318.3
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: EP 01870002.1
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: EP 01870003.9
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 732
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 414
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-451-467A-414
```

Query Match 70.2%; Score 59; DB 16; Length 171;
Best Local Similarity 58.3%; Pred. No. 1.2;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHHHHHHHHH 12
|:|:|:|:|:|:|

Db 100 HHHHHHHHHH 111

```
RESULT 58
US-10-029-386-34005
; Sequence 34005, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34286
; SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34005
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC023112.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.39
US-10-029-386-34005
```

Query Match 70.2%; Score 59; DB 14; Length 186;
Best Local Similarity 58.3%; Pred. No. 1.3;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHHHHHHHHH 12
|:|:|:|:|:|:|

Db 151 HHHHHHHHHH 162

```
RESULT 59
US-10-424-599-155301
; Sequence 155301, Application US/10424599
```

```

; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 155301
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(217)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: clone ID: PAT_MRT3847_111258C.1.pep
; US-10-424-599-155301

Query Match          70.2%; Score 59; DB 15; Length 217;
Best Local Similarity 58.3%; Pred. No. 1.5;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY      1 HNNHNNHNNH 12
Db      182 HHHHHHHHHH 193

RESULT 60
; US-10-282-122A-63467
; Sequence 63467, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308

```

```

; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63467
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Mycoplasma genitalium
; US-10-282-122A-63467

Query Match          70.2%; Score 59; DB 15; Length 281;
Best Local Similarity 72.7%; Pred. No. 1.8;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY      1 HNNHNNHNNH 11
Db      146 HDHNNHNNHH 156

RESULT 61
; US-09-864-761-37944
; Sequence 37944, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37944
; LENGTH: 292

```



```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005414.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN HELI00, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 11
US-09-864-761-37944
```

```

Query Match          70.2%; Score 59; DB 9; Length 292;
Best Local Similarity 58.3%; Pred. No. 1.9;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 HNNNNNNNNNN 12
        |:|:|:|:|:|:|
Db      56 HHHHHHHHHH 67
```

```

RESULT 62
US-11-097-143-10602
; Sequence 10602, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10602
; LENGTH: 299
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-10602
```

```

Query Match          70.2%; Score 59; DB 20; Length 299;
Best Local Similarity 58.3%; Pred. No. 1.9;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 HNNNNNNNNNN 12
        |:|:|:|:|:|:|
Db      115 HHHHHHHHHH 126
```

```

RESULT 63
US-11-097-143-8499
```

```

; Sequence 8499, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8499
; LENGTH: 374
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-8499
```

```

Query Match          70.2%; Score 59; DB 20; Length 374;
Best Local Similarity 58.3%; Pred. No. 2.4;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 HNNNNNNNNNN 12
        |:|:|:|:|:|:|
Db      164 HSHHSHDHNN 175
```

```

RESULT 64
US-10-741-849-7304
; Sequence 7304, Application US/10741849
; Publication No. US20050019931A1
; GENERAL INFORMATION:
; APPLICANT: Roemer, Terry
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Nucleic Acids Encoding Anti-fungal Drug Targets and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 10182-023-999
; CURRENT APPLICATION NUMBER: US/10/741,849
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: US 60/434,832
; PRIOR FILING DATE: 2002-12-19
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 7304
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-741-849-7304
```

```

Query Match          70.2%; Score 59; DB 17; Length 467;
Best Local Similarity 81.8%; Pred. No. 2.9;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```

QY      2 NNNNNNNNNN 12
```

Db 78 NTNNNNNNNN 88

RESULT 65

```

US-11-097-143-4347
; Sequence 4347, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4347
; LENGTH: 537
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-4347

```

Query Match Best Local Similarity 70.2%; Score 59; DB 20; Length 537;

Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12

Db 355 HSHSHHHHHN 366

RESULT 66

```

US-11-097-143-20661
; Sequence 20661, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28

```

PRIOR APPLICATION NUMBER: 60/175,693

PRIOR FILING DATE: 2000-01-12

PRIOR APPLICATION NUMBER: 60/184,831

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: 60/191,637

PRIOR FILING DATE: 2000-03-23

NUMBER OF SEQ ID NOS: 43008

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 20661

LENGTH: 989

TYPE: PRT

ORGANISM: DROSOPHILA

US-11-097-143-20661

Query Match Best Local Similarity 70.2%; Score 59; DB 20; Length 989;

Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12

Db 188 HNNNNNNNNNN 199

RESULT 67

```

US-11-097-143-31017
; Sequence 31017, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31017
; LENGTH: 1300
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-31017

```

Query Match Best Local Similarity 70.2%; Score 59; DB 20; Length 1300;

Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12

Db 544 HNNNNNNNNNN 555

RESULT 68

```

US-10-450-763-40872
; Sequence 40872, Application US/10450763
; Publication No. US20050196754A1

```

```
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc
/ TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
/ FILE REFERENCE: 790CIP3/US
/ CURRENT APPLICATION NUMBER: US/10/450,763
/ PRIOR FILING DATE: 2003-06-11
/ PRIOR APPLICATION NUMBER: PCT/US01/08631
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: 09/540,217
/ PRIOR FILING DATE: 2000-03-31
/ PRIOR APPLICATION NUMBER: 09/649,167
/ PRIOR FILING DATE: 2000-08-23
/ NUMBER OF SEQ ID NOS: 60736
/ SOFTWARE: Custom
/ SEQ ID NO 40872
/ LENGTH: 51
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-450-763-40872
```

```
Query Match          69.0%; Score 58; DB 18; Length 51;
Best Local Similarity 63.6%; Pred. No. 0.53;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 HNNNNNNNNH 11
       |:|:|:|:|:|
Db      7 HHHHHHHHHH 17
```

```
RESULT 69
US-10-450-763-57203
/ Sequence 57203, Application US/10450763
/ Publication No. US20050196754A1
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc
/ TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
/ FILE REFERENCE: 790CIP3/US
/ CURRENT APPLICATION NUMBER: US/10/450,763
/ CURRENT FILING DATE: 2003-06-11
/ PRIOR APPLICATION NUMBER: PCT/US01/08631
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: 09/540,217
/ PRIOR FILING DATE: 2000-03-31
/ PRIOR APPLICATION NUMBER: 09/649,167
/ PRIOR FILING DATE: 2000-08-23
/ NUMBER OF SEQ ID NOS: 60736
/ SOFTWARE: Custom
/ SEQ ID NO 57203
/ LENGTH: 117
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-450-763-57203
```

```
Query Match          69.0%; Score 58; DB 18; Length 117;
Best Local Similarity 58.3%; Pred. No. 1.1;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 HNNNNNNNNH 12
       |:|:|:|:|:|
Db      59 YNNSHHHHHHH 70
```

```
RESULT 70
US-09-864-761-41679
/ Sequence 41679, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
```

```
/ FILE REFERENCE: Aecmlca-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 41679
/ LENGTH: 156
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AP001347.1
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
US-09-864-761-41679
```

```
Query Match          69.0%; Score 58; DB 9; Length 156;
Best Local Similarity 58.3%; Pred. No. 1.5;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 HNNNNNNNNH 12
       |:|:|:|:|:|
Db      100 HHHHHHHHHH 111
```

```
RESULT 71
US-10-114-893-42
/ Sequence 42, Application US/10114893
/ Publication No. US20020193567A1
/ GENERAL INFORMATION:
/ APPLICANT: Jacobs, Kenneth
/ APPLICANT: McCoy, John M.
/ APPLICANT: Lavallee, Edward R.
/ APPLICANT: Collins-Racie, Lisa A.
/ APPLICANT: Evans, Cheryl
```

APPLICANT: Werberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Bowman, Michael R.
APPLICANT: Spaulding, Vikki
APPLICANT: Carlin-Duckett, McKeough
APPLICANT: Kelleher, Kerry S.
TITLE OF INVENTION: GENETICS INSTITUTE, INC.
FILE REFERENCE: GI 6000-10A
CURRENT APPLICATION NUMBER: US/10-114,893
CURRENT FILING DATE: 2002-04-02
EARLIER FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 321
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 42
LENGTH: 164
TYPE: PRT
ORGANISM: Homo sapiens
US-10-114-893-42

Query Match 69.0%; Score 58; DB 13; Length 164;
Best Local Similarity 58.3%; Pred. No. 1.5;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNHNNHNNH 12
|:|:|:|:|:|:
DB 18 YNHSNNHHHHH 29

RESULT 72
US-10-450-763-50868
Sequence 50868, Application US/10450763
Publication No. US20050196754A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
FILE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
TITLE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 50868
LENGTH: 324
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (48)..(72)
OTHER INFORMATION: HMW KININOGEN SIGNATURE domain identified by eMATRIX,
FILE REFERENCE: accession number PR00334B, p-value=9.242e-11, raw score of 8.69
NAME/KEY: misc feature
LOCATION: (1)..(324)
OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-50868

Query Match 69.0%; Score 58; DB 18; Length 324;
Best Local Similarity 58.3%; Pred. No. 2.8;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNHNNHNNH 12
|:|:|:|:|:|:
DB 24 HHHHHHHNNHH 35

RESULT 73

US-11-097-143-36393
Sequence 36393, Application US/11097143
Publication No. US20050208558A1
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
APPLICANT: et al.
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
FILE REFERENCE: CL000728
CURRENT APPLICATION NUMBER: US/11/097,143
CURRENT FILING DATE: 2005-04-04
PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/160,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/164,769
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/173,383
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: 60/175,693
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: 60/184,831
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637
PRIOR FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 43008
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 36393
LENGTH: 368
TYPE: PRT
ORGANISM: DROSOPHILA
US-11-097-143-36393

Query Match 69.0%; Score 58; DB 20; Length 368;
Best Local Similarity 58.3%; Pred. No. 3.1;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 HNNHNNHNNH 12
|:|:|:|:|:|:
DB 260 HSHTHSHSHSHN 271

RESULT 74
US-10-029-386-33892
Sequence 33892, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Hanzel, David R.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: ABOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 33892
LENGTH: 378
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MAP TO AP001751.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.6
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
US-10-029-386-33892

Query Match 69.0%; Score 58; DB 14; Length 378;
Best Local Similarity 58.3%; Pred. No. 3.2;

Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 HNNHNNHNNH 12
|:|:|:|:|:|:
Db 327 HHHHHHHHHH 338

RESULT 75
US-11-097-143-36015
; Sequence 36015, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36015
; LENGTH: 499
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-36015

Query Match 69.0%; Score 58; DB 20; Length 499;
Best Local Similarity 80.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 10
|:|:|:|:|:|:
Db 25 HGHGHNNHNN 34

RESULT 76
US-11-097-143-2451
; Sequence 2451, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769

; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2451
; LENGTH: 649
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-2451

Query Match 69.0%; Score 58; DB 20; Length 649;
Best Local Similarity 72.7%; Pred. No. 5.2;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 11
|:|:|:|:|:|:
Db 618 NSHNNHNNH 628

RESULT 77
US-09-890-688-130
; Sequence 130, Application US/09890688
; Publication No. US20030144475A1
; GENERAL INFORMATION:
; APPLICANT: Selski KATO
; APPLICANT: Chikashi EGUCHI
; APPLICANT: Mihoro SAKETI
; TITLE OF INVENTION: Human Proteins and cDNAs thereof
; FILE REFERENCE: 2001-1102A/WMC/00653
; CURRENT APPLICATION NUMBER: US/09/890,688
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: JP 11-346863
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: JP 11-34684
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: JP 2000-31062
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: JP 2000-34091
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: JP 2000-34090
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: JP 2000-35829
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: JP 2000-35899
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: JP 2000-71161
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: JP 2000-160851
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 130
; LENGTH: 695
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-890-688-130

Query Match 69.0%; Score 58; DB 10; Length 695;
Best Local Similarity 58.3%; Pred. No. 5.6;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNHNNH 12
|:|:|:|:|:|:
Db 39 YNHSNNHHHH 50

RESULT 78
 US-11-097-143-23127
 ; Sequence 23127, Application US/11097143
 ; Publication No. US20050208558A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Venter, J. Craig
 ; APPLICANT: et al.
 ; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
 ; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
 ; FILE REFERENCE: CL000728
 ; CURRENT APPLICATION NUMBER: US/11/097,143
 ; PRIORITY FILING DATE: 2005-04-04
 ; PRIOR APPLICATION NUMBER: 60/157,832
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: 60/160,191
 ; PRIOR FILING DATE: 1999-10-19
 ; PRIOR APPLICATION NUMBER: 60/161,932
 ; PRIOR FILING DATE: 1999-10-28
 ; PRIOR APPLICATION NUMBER: 60/164,769
 ; PRIOR FILING DATE: 1999-11-12
 ; PRIOR APPLICATION NUMBER: 60/173,383
 ; PRIOR FILING DATE: 1999-12-28
 ; PRIOR APPLICATION NUMBER: 60/175,693
 ; PRIOR FILING DATE: 2000-01-12
 ; PRIOR APPLICATION NUMBER: 60/184,831
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: 60/191,637
 ; PRIOR FILING DATE: 2000-03-23
 ; NUMBER OF SEQ ID NOS: 43008
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 23127
 ; LENGTH: 749
 ; TYPE: PRT
 ; ORGANISM: DROSOPHILA
 US-11-097-143-23127

Query Match 69.0%; Score 58; DB 20; Length 749;
 Best Local Similarity 58.3%; Pred. No. 5.9;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HNNNNNNNNNN 12
 |||:|:|:|:
 Db 515 HNNSHDHSQHS 526

RESULT 79
 US-10-754-896-2
 ; Sequence 2, Application US/10754896
 ; Publication No. US20040197808A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hall, Ronald J.
 ; APPLICANT: Hanan, Garry N.
 ; TITLE OF INVENTION: Genetic Sequences Encoding Steroid and Juvenile Hormone
 ; TITLE OF INVENTION: Receptor Polypeptides and Insecticidal Modalities
 ; TITLE OF INVENTION: Therefor
 ; FILE REFERENCE: 53-99
 ; CURRENT APPLICATION NUMBER: US/10/754,896
 ; CURRENT FILING DATE: 2004-01-08
 ; PRIOR APPLICATION NUMBER: WO 99/00033
 ; PRIOR FILING DATE: 1999-01-15
 ; PRIOR APPLICATION NUMBER: AU PPI356/98
 ; PRIOR FILING DATE: 1998-01-15
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 757
 ; TYPE: PRT
 ; ORGANISM: Lucilia cuprina
 US-10-754-896-2

Query Match 69.0%; Score 58; DB 16; Length 757;
 Best Local Similarity 66.7%; Pred. No. 6;

Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HNNNNNNNNNN 12
 :|||:|:|:|:
 Db 155 NNNNNSHNNLHH 166

RESULT 80
 US-11-097-143-23004
 ; Sequence 23004, Application US/11097143
 ; Publication No. US20050208558A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Venter, J. Craig
 ; APPLICANT: et al.
 ; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
 ; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
 ; FILE REFERENCE: CL000728
 ; CURRENT APPLICATION NUMBER: US/11/097,143
 ; PRIORITY FILING DATE: 2005-04-04
 ; PRIOR APPLICATION NUMBER: 60/157,832
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: 60/160,191
 ; PRIOR FILING DATE: 1999-10-19
 ; PRIOR APPLICATION NUMBER: 60/161,932
 ; PRIOR FILING DATE: 1999-10-28
 ; PRIOR APPLICATION NUMBER: 60/164,769
 ; PRIOR FILING DATE: 1999-11-12
 ; PRIOR APPLICATION NUMBER: 60/173,383
 ; PRIOR FILING DATE: 1999-12-28
 ; PRIOR APPLICATION NUMBER: 60/175,693
 ; PRIOR FILING DATE: 2000-01-12
 ; PRIOR APPLICATION NUMBER: 60/184,831
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: 60/191,637
 ; NUMBER OF SEQ ID NOS: 43008
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 23004
 ; LENGTH: 789
 ; TYPE: PRT
 ; ORGANISM: DROSOPHILA
 US-11-097-143-23004

Query Match 69.0%; Score 58; DB 20; Length 789;
 Best Local Similarity 58.3%; Pred. No. 6.2;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNNNNNNNNN 12
 |||:|:|:|:
 Db 555 HNNSHDHSQHS 566

RESULT 81
 US-10-424-599-175849
 ; Sequence 175849, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 175849
 ; LENGTH: 178
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_12980C.1.pep
US-10-424-599-175849

Query Match 68.5%; Score 57.5; DB 15; Length 178;
Best Local Similarity 75.0%; Pred. No. 1.9;
Matches 9; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 HNNHNH-NHNNH 11
|:|:|:|:|:|:
Db 3 HNNHHRHNNH 14

RESULT 82

US-10-424-599-193590
Sequence 193590, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21 (53223) B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 193590
LENGTH: 188
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_16836C.1.pep
US-10-424-599-193590

Query Match 67.9%; Score 57; DB 15; Length 188;
Best Local Similarity 58.3%; Pred. No. 2.3;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNH-NHNNH 12
|:|:|:|:|:|:
Db 176 HNNHNNHNNH 187

RESULT 83

US-10-424-599-169851
Sequence 169851, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21 (53223) B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 169851
LENGTH: 193
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_12438C.1.pep
US-10-424-599-169851

Query Match 67.9%; Score 57; DB 15; Length 193;
Best Local Similarity 72.7%; Pred. No. 2.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNH-NHNNH 11
|:|:|:|:|:|:

Db 174 HSEPTNNHNNH 184

RESULT 84

US-10-450-763-44758
Sequence 44758, Application US/10450763
Publication No. US20050196754A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 44758
LENGTH: 227
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (63)..(77)
OTHER INFORMATION: PROTEIN GLYCOPROTEIN PRECURSOR RE domain identified by
OTHER INFORMATION: EMATRIX, accession number PD00306A, p-value=3.250e-10, raw score
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(227)
OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-44758

Query Match 67.9%; Score 57; DB 18; Length 227;
Best Local Similarity 58.3%; Pred. No. 2.7;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 HNNHNH-NHNNH 12
|:|:|:|:|:|:
Db 150 HNNHNNHNNH 161

RESULT 85

US-09-934-455-176
Sequence 176, Application US/09934455
Publication No. US20030121070A1
GENERAL INFORMATION:
APPLICANT: Adam, Luc
APPLICANT: Creelman, Robert
APPLICANT: Dubell, Arnold
APPLICANT: Heard, Jacqueline
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Keddle, James
APPLICANT: Pilgrim, Marsha
APPLICANT: Ratcliffe, Oliver
APPLICANT: Reuber, Lynne
APPLICANT: Riechmann, Jose Luis
APPLICANT: Yu, Guo-Liang
APPLICANT: Pineda, Omatia
TITLE OF INVENTION: Genes for Modifying Plant Traits IV
FILE REFERENCE: MBI-0025
CURRENT APPLICATION NUMBER: US/09/934,455
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/227439
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: MBI-0022
PRIOR FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: MBI-0023
PRIOR FILING DATE: 2001-04-17

```

; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 176
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-934-455-176
```

```

Query Match      67.9%; Score 57; DB 10; Length 285;
Best Local Similarity 58.3%; Pred. No. 3.4;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 HHHHHHHHHN 12
        ::|::|::|::|:
DB      39 NHHHHHHHHH 50
```

```

RESULT 86
US-10-278-173-16
; Sequence 16, Application US/10278173
; Publication NO. US20030061637A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Pineda, Omaira
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Keddie, James
; APPLICANT: Heard, Jacqueline
; APPLICANT: Reuber, Lynne
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; TITLE OF INVENTION: POLYNUCLEOTIDES FOR ROOT TRAIT ALTERATION
; FILE REFERENCE: MBI-009
; CURRENT APPLICATION NUMBER: US/10/278,173
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 16
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1075
US-10-278-173-16
```

```

Query Match      67.9%; Score 57; DB 14; Length 285;
Best Local Similarity 58.3%; Pred. No. 3.4;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 HHHHHHHHHN 12
        ::|::|::|::|:
DB      39 NHHHHHHHHH 50
```

```

RESULT 87
US-10-225-066A-56
; Sequence 56, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J
; APPLICANT: Dubell, Arnold T
; APPLICANT: Heard, Jacqueline E
```

```

; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Creelman, Robert A
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Broun, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI0036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 56
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-56
```

```

Query Match      67.9%; Score 57; DB 15; Length 285;
Best Local Similarity 58.3%; Pred. No. 3.4;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 HHHHHHHHHN 12
        ::|::|::|::|:
DB      39 NHHHHHHHHH 50
```

```

RESULT 88
US-10-374-780A-226
; Sequence 226, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; PRIOR FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
```



```

; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 226
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1075 (conserved domain in AA coordinates: 78-85)
US-10-374-780A-226

Query Match      67.9%; Score 57; DB 15; Length 285;
Best Local Similarity 58.3%; Pred. No. 3.4;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1 HNNNNNNNNNN 12
DB      39 NHHHHHHNNHH 50

RESULT 89
US-10-412-699B-726
; Sequence 726, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline F.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omalra
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: Dubeil, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
```

```

; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 726
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1075
US-10-412-699B-726

Query Match      67.9%; Score 57; DB 15; Length 285;
Best Local Similarity 58.3%; Pred. No. 3.4;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1 HNNNNNNNNNN 12
DB      39 NHHHHHHNNHH 50

RESULT 90
US-10-669-824-52
; Sequence 52, Application US/10669824
; Publication No. US20040128712A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Guterson, Neal
; APPLICANT: Hempel, Frederick
; APPLICANT: Kumimoto, Roderick W
; APPLICANT: Keddle, James S
; APPLICANT: Sherman, Bradley K
; TITLE OF INVENTION: METHODS FOR MODIFYING PLANT BIOMASS AND TOLERANCE TO ABIOTIC
; FILE REFERENCE: MBI-0034CIP
; CURRENT APPLICATION NUMBER: US/10/669,824
; CURRENT FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: US 10/374,780
; PRIOR FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: US 60/227,439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: US 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: US 60/135,134
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/125,814
; PRIOR FILING DATE: 1999-03-23
; Remaining Prior Application data removed - See File wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1075 polypeptide
US-10-669-824-52

Query Match      67.9%; Score 57; DB 16; Length 285;
Best Local Similarity 58.3%; Pred. No. 3.4;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 HNNNNNNNNN 12
: : : : :
Db 39 NHHHHHHNNHH 50

RESULT 91

US-10-870-198-52
; Sequence 52, Application US/10870198
; Publication No. US20050097638A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: GUTTERSON, Neal I
; APPLICANT: HEMPEL, Frederick D
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: KEDDIE, James S
; APPLICANT: SHERMAN, Bradley K
; APPLICANT: LIBBY, Jeffrey M
; TITLE OF INVENTION: Transcriptional Regulation of Plant Biomass and Abiotic Stresses
; FILE REFERENCE: MBI-0068CIP
; CURRENT APPLICATION NUMBER: US/10/870,198
; PRIOR FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/565,948
; PRIOR FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 10/669,824
; PRIOR FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: 10/374,780
; PRIOR FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227,439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1075 polypeptide
US-10-870-198-52

Query Match 67.9%; Score 57; DB 17; Length 285;
Best Local Similarity 58.3%; Pred. No. 3.4;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNNNNNNNN 12
: : : : :
Db 39 NHHHHHHNNHH 50

RESULT 92
US-10-225-066A-56
; Sequence 56, Application US/10225066A
; Publication No. US20050160493A9
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver

; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROON, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI0036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-56

Query Match 67.9%; Score 57; DB 18; Length 285;
Best Local Similarity 58.3%; Pred. No. 3.4;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 1 HNNNNNNNNN 12
: : : : :
Db 39 NHHHHHHNNHH 50

RESULT 93
US-11-097-143-5616
; Sequence 5616, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5616
LENGTH: 485
TYPE: PRT
ORGANISM: DROSOPHILA
US-11-097-143-5616

Query Match 67.9%; Score 57; DB 20; Length 485;
Best Local Similarity 66.7%; Pred. No. 5.4;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HNNHNNNNNN 12
|:|:|:|:|:|:|
DB 342 HNNHNNNNNN 353

RESULT 94
US-11-013-314-38
Sequence 38, Application US/11013314
Publication No. US20050166269A1
GENERAL INFORMATION:
APPLICANT: TONONI, GIULIO
TITLE OF INVENTION: SLEEP GENES IN DROSOPHILA AND THEIR USE FOR THE
TITLE OF INVENTION: SCREENING, DIAGNOSIS AND THERAPY OF SLEEP DISORDERS
FILE REFERENCE: WARP:010US
CURRENT APPLICATION NUMBER: US/11/013,314
CURRENT FILING DATE: 2004-12-15
PRIOR APPLICATION NUMBER: 60/563,858
PRIOR FILING DATE: 2004-04-20
PRIOR APPLICATION NUMBER: 60/529,536
PRIOR FILING DATE: 2003-12-15
NUMBER OF SEQ ID NOS: 55
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 38
LENGTH: 660
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-11-013-314-38

Query Match 67.9%; Score 57; DB 20; Length 660;
Best Local Similarity 56.2%; Pred. No. 7.1;
Matches 9; Conservative 3; Mismatches 0; Indels 4; Gaps 1;

QY 1 HNNHNNNNNN 12
|:|:|:|:|:|:|
DB 56 HHHHPLNHHNNNN 71

RESULT 95
US-11-097-143-14205
Sequence 14205, Application US/11097143
Publication No. US20050208558A1
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
APPLICANT: et al.
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
TITLE OF INVENTION: DROSOPHILA GENES.
FILE REFERENCE: CL000728
CURRENT APPLICATION NUMBER: US/11/097,143
CURRENT FILING DATE: 2005-04-04
PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/160,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/164,769
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/173,383
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: 60/175,693
PRIOR FILING DATE: 2000-01-12

PRIOR APPLICATION NUMBER: 60/184,831
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637
PRIOR FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 43008
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14205
LENGTH: 660
TYPE: PRT
ORGANISM: DROSOPHILA
US-11-097-143-14205

Query Match 67.9%; Score 57; DB 20; Length 660;
Best Local Similarity 56.2%; Pred. No. 7.1;
Matches 9; Conservative 3; Mismatches 0; Indels 4; Gaps 1;

QY 1 HNNHNNNNNN 12
|:|:|:|:|:|:|
DB 56 HHHHPLNHHNNNN 71

RESULT 96
US-10-741-849-7248
Sequence 7248, Application US/10741849
Publication No. US20050019931A1
GENERAL INFORMATION:
APPLICANT: Roemer, Terry
APPLICANT: Uiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Nucleic Acids Encoding Anti-fungal Drug Targets and Methods of
TITLE OF INVENTION: Use
FILE REFERENCE: 10182-023-999
CURRENT APPLICATION NUMBER: US/10/741,849
CURRENT FILING DATE: 2003-12-19
PRIOR APPLICATION NUMBER: US 60/434,832
PRIOR FILING DATE: 2002-12-19
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.2
SEQ ID NO 7248
LENGTH: 1146
TYPE: PRT
ORGANISM: Candida albicans
US-10-741-849-7248

Query Match 67.9%; Score 57; DB 17; Length 1146;
Best Local Similarity 80.0%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HNNHNNNNNN 10
|:|:|:|:|:|:|
DB 1062 NNHNNNNNN 1071

RESULT 97
US-11-097-143-38049
Sequence 38049, Application US/11097143
Publication No. US20050208558A1
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
APPLICANT: et al.
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
TITLE OF INVENTION: DROSOPHILA GENES.
FILE REFERENCE: CL000728
CURRENT APPLICATION NUMBER: US/11/097,143
CURRENT FILING DATE: 2005-04-04
PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/160,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28

```

; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38049
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-38049

```

```

Query Match      67.9%; Score 57; DB 20; Length 1161;
Best Local Similarity 58.3%; Pred. No. 12;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 HNNNNNNNNNN 12
       727 HHHSHHHGYNHN 738

```

```

RESULT 98
US-10-450-763-38565
; Sequence 38565, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 38565
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (7)..(31)
; OTHER INFORMATION: HMW KININOGEN SIGNATURE domain identified by eMATRIX.
; OTHER INFORMATION: accession number PR00334B, p-value=2.980e-10, raw score of 8.69
; NAME/KEY: misc_feature
; LOCATION: (1)...(204)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-38565

```

```

Query Match      66.7%; Score 56; DB 18; Length 204;
Best Local Similarity 58.3%; Pred. No. 3;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY      1 HNNNNNNNNNN 12
       10 HHHHHHHHHHH 21

```

```

RESULT 99
US-10-451-467A-296
; Sequence 296, Application US/10451467A
; Publication No. US20040161840A1

```

```

; GENERAL INFORMATION:
; APPLICANT: CONTRERAS, ROLAND HENRI
; APPLICANT: EBERHARDT, INES
; APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS
; APPLICANT: REEKMAN, RIEKA JOSEPHINA
; TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN
; FILE REFERENCE: YAB-1667
; CURRENT APPLICATION NUMBER: US/10/451,467A
; CURRENT FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: EP 00870318.3
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: EP 01870002.1
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: EP 01870003.9
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 732
; SOFTWARE: Patencin version 3.1
; SEQ ID NO 296
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-451-467A-296

```

```

Query Match      66.7%; Score 56; DB 16; Length 212;
Best Local Similarity 56.7%; Pred. No. 3.5;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 HNNNNNNNNNN 12
       90 NNNNNNNNNNN 101

```

```

RESULT 100
US-10-087-887-76
; Sequence 76, Application US/10087887
; Publication No. US20030198957A1
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Conley, Pamela B.
; APPLICANT: Yang, Ruey-Bing
; APPLICANT: Hart, Matthew
; APPLICANT: Tomlinson, James E.
; APPLICANT: Topper, James N.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Leach, Martin D.
; APPLICANT: Zethusen, Bryan D.
; APPLICANT: Komuves, Laszlo
; APPLICANT: Padigattu, Muralidhara
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-285
; CURRENT APPLICATION NUMBER: US/10/087,887
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/273,049
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/279,883
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/277,791
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/281,248
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/282,864
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,537
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/282,867
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 76
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-10-087-887-76

Query Match 66.7%; Score 56; DB 14; Length 407;
Best Local Similarity 63.6%; Pred. No. 6.2;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNNNNNNNNN 12
:|:|:|:|:|:|
Db 305 HHHHHHHHHN 315

Search completed: October 4, 2005, 11:44:43
Job time : 1243 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 4, 2005, 11:44:51 ; Search time 1815 Seconds

(without alignments)
320.365 Million cell updates/sec

Title: US-09-858-332g-15

Sequence: 1 HNNHNNHNNHNN 12

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:
-MODE=frame_plus_p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US09858332/rnat 04102005 122256 4293/app.query.fasta_1.199
-DB=GenEmbl -QFMT=fastap -SUPFI=sp2n.rge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdt -LIST=100
-DOCALLIGN=200 -THR.SCORE=Pct -THR.MAX=100 -THR.MIN=0 -ALIGN=100 -MODE=LOCAL
-OUTFMT=Pct -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09858332@cgn 1.1 5600 @rnat 04102005 122256 4293 -NCPU=6 -ICPU=3
-NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_da.*
2: gb_hlg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description
1	84	100.0	42	6	AX594160	AX594160	Sequence
2	84	100.0	100	6	AX594158	AX594158	Sequence
3	84	100.0	115	4	AY559840	Bos tauru	
4	84	100.0	387	4	AF097207	Bos tauru	

5	84	100.0	390	4	AF097209	Bos tauru	
6	84	100.0	427	3	PF430088	Plasmodi	
7	84	100.0	478	10	BTU49779	Bos tauru	
8	84	100.0	519	10	AF187065	Rattus no	
9	84	100.0	887	10	BC058503	Rattus no	
10	84	100.0	994	11	CNS065H9	R3 end of	
11	84	100.0	1132	8	AY065142	Arabidops	
12	84	100.0	1897	6	AK107456	Oryza sat	
13	84	100.0	1704	8	IC4914	Sequence 1	
14	84	100.0	1897	8	YSCOP11	Saccharomyc	
15	84	100.0	1898	6	AX536592	Sequence	
16	84	100.0	1949	10	BC029741	Mus muscu	
17	84	100.0	2011	10	AE039933	Mus muscu	
18	84	100.0	2152	10	BC083830	Rattus no	
19	84	100.0	2270	6	AX461526	Sequence	
20	84	100.0	2396	6	AF020407	Dictyoste	
21	84	100.0	4938	6	AX594148	Sequence	
22	84	100.0	6525	6	AX594150	Sequence	
23	84	100.0	7487	6	AX594151	Sequence	
24	84	100.0	7801	3	AY160101	Dictyoste	
25	84	100.0	9024	3	AF411930	Plasmodi	
26	84	100.0	9044	3	AF411933	Plasmodi	
27	84	100.0	9049	3	AF411929	Plasmodi	
28	84	100.0	29276	8	YSC9433	Saccharomyc	
29	84	100.0	32083	8	AF309805	Pneumococ	
30	84	100.0	64707	3	AC115607	Dictyoste	
31	84	100.0	65263	2	AC113185	Mus muscu	
32	84	100.0	75741	8	AC149204	Medicago	
33	84	100.0	79776	2	AC139431	Homo sapi	
34	84	100.0	83029	8	ATT805	Arabidops	
35	84	100.0	101256	8	CNS0771M	DNA cent	
36	84	100.0	101752	8	AP006378	Lotus cor	
37	84	100.0	110000	2	AC109729_3	Continuati	
38	84	100.0	110000	2	AC120234_1	Continuati	
39	84	100.0	110000	2	AC142303_3	Continuati	
40	84	100.0	110000	8	CR382132_16	Continuati	
41	84	100.0	110000	8	AE016820_05	Continuati	
42	84	100.0	112057	2	AP004065	Oryza sat	
43	84	100.0	125623	3	AC115599	Dictyoste	
44	84	100.0	125958	3	AC115592	Dictyoste	
45	84	100.0	127902	5	AP005406	Oryza sat	
46	84	100.0	144092	8	BX465845	Zebrafish	
47	84	100.0	153751	5	AC116551	Dictyoste	
48	84	100.0	157087	2	AC142155	Rattus no	
49	84	100.0	157193	9	AL353515	Human DNA	
50	84	100.0	161701	2	AC117110	Rattus no	
51	84	100.0	161965	10	AC148008	Rattus no	
52	84	100.0	163087	2	AC139383	Rattus no	
53	84	100.0	170533	9	AP003356	Homo sapi	
54	84	100.0	176733	2	AC068001	Homo sapi	
55	84	100.0	179415	10	AC138578	Mus muscu	
56	84	100.0	182870	3	AC116960	Dictyoste	
57	84	100.0	182871	3	AC117176	Dictyoste	
58	84	100.0	184202	2	AC116475	Mus muscu	
59	84	100.0	184541	9	AC011405	Homo sapi	
60	84	100.0	188388	2	AC131664	Mus muscu	
61	84	100.0	188453	10	AC123741	Mus muscu	
62	84	100.0	189105	2	CR450684	Danio rer	
63	84	100.0	189130	10	AC132110	Mus muscu	
64	84	100.0	189505	10	AC134540	Mus muscu	
65	84	100.0	191308	9	AC034243	Homo sapi	
66	84	100.0	194916	8	ATCHRIV55	Arabidops	
67	84	100.0	196247	8	ATCHRIV56	Arabidops	
68	84	100.0	198321	10	AC104863	Mus muscu	
69	84	100.0	198341	5	AC122454	Mus muscu	
70	84	100.0	201932	5	BX897729	Zebrafish	
71	84	100.0	204940	9	AC069007	Homo sapi	
72	84	100.0	207611	9	AC091132	Homo sapi	
73	84	100.0	210632	10	AC120797	Mus muscu	
74	84	100.0	218133	9	AC074051	Homo sapi	
75	84	100.0	220006	2	AC084068	Mus muscu	
76	84	100.0	226013	2	AC073680	Mus muscu	
77	84	100.0	227081	10	AC119847	Mus muscu	

/note="recombination signal sequence"
107.115
/note="recombination signal sequence"

ORIGIN

Alignment Scores:

Pred. No.:	0.000907	Length:	115
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0

DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x AY559840 (1-115)

Qy 1 HisAenHisAenHisAenHisAenHisAen 12
|||||
80 CATACCATTAACCATTAACCATTAACCATTAAC 45

RESULT 4
AF097207/c
LOCUS AF097207 387 bp mRNA linear MAM 17-DEC-2003
DEFINITION Bos taurus clone BF3H11 immunoglobulin heavy chain variable region
ACCESSION AF097207
VERSION AF097207.1 GI:3834654
KEYWORDS
SOURCE Bos taurus (cow)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

REFERENCE
AUTHORS Saini,S.S. and Kaushik,A.
TITLE Extensive CDR3H length heterogeneity exists in bovine foetal VDJ rearrangements
JOURNAL Scand. J. Immunol. 55 (2), 140-148 (2002)
MEDLINE 21896439
PubMed 11896930

AUTHORS Saini,S.S., Allore,B., Jacobs,R. and Kaushik,A.
TITLE Direct Submision
JOURNAL Submitted (07-OCT-1998) Pathobiology, University of Guelph, Guelph, ON N1G 2W1, Canada

FEATURES
source Location/Qualifiers
1..387
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="BF3H11"
/cell_type="mouse x Bovine heterohybridoma"
/dev_stage="fetal"
/note="PCR product"
<1..>387
/note="VDJ region"
/codon_start=1
/product="immunoglobulin heavy chain variable region"
/protein_id="AAC71042.1"
/db_xref="GI:3834654"
/translation="QVQLRSGPSLVKPSQTLSTCTVSGFSLSSNGVGVWRQAPGKA LEWVGIDNDGTYNPAKSRSLSTKDNKSQVSLSSVTPEDTATYYCARDCYSG YGVCYGYGYDYVDAMGQGLVTVSS"

ORIGIN

Alignment Scores:

Pred. No.:	0.0032	Length:	387
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0

DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x AF097207 (1-387)

Qy 1 HisAenHisAenHisAenHisAenHisAen 12
|||||
340 CATACCATTAACCATTAACCATTAACCATTAAC 305

RESULT 5
AF097209/c
LOCUS AF097209 390 bp mRNA linear MAM 17-DEC-2003
DEFINITION Bos taurus clone BFSF10 immunoglobulin heavy chain variable region
ACCESSION AF097209
VERSION AF097209.1 GI:3834658
KEYWORDS
SOURCE Bos taurus (cow)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

REFERENCE
AUTHORS Saini,S.S. and Kaushik,A.
TITLE Extensive CDR3H length heterogeneity exists in bovine foetal VDJ rearrangements
JOURNAL Scand. J. Immunol. 55 (2), 140-148 (2002)
MEDLINE 21896439
PubMed 11896930

AUTHORS Saini,S.S., Allore,B., Jacobs,R. and Kaushik,A.
TITLE Direct Submision
JOURNAL Submitted (07-OCT-1998) Pathobiology, University of Guelph, Guelph, ON N1G 2W1, Canada

FEATURES
source Location/Qualifiers
1..390
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="BFSF10"
/cell_type="mouse x Bovine heterohybridoma"
/dev_stage="fetal"
/note="PCR product"
<1..>390
/note="VDJ region"
/codon_start=1
/product="immunoglobulin heavy chain variable region"
/protein_id="AAC71044.1"
/db_xref="GI:3834659"
/translation="QVQLRSGPSLVKPSQTLSTCTVSGFSLSSNGVGVWRQAPGKA LEWVGITSGSGTGNPAKSRSLSTKDNKSQVSLSSVTPEDTATYYCARDCYSG YGVCYGYGYDYVDAMGQGLVTVSS"

ORIGIN

Alignment Scores:

Pred. No.:	0.00323	Length:	390
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0

DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x AF097209 (1-390)

Qy 1 HisAenHisAenHisAenHisAenHisAen 12
|||||
340 CATACCATTAACCATTAACCATTAACCATTAAC 305

RESULT 6
PFA430088
LOCUS PFA430088 427 bp DNA linear INV 07-FEB-2002
DEFINITION Plasmodium falciparum partial rhl gene, isolate FCBI.
ACCESSION AJ430088
VERSION AJ430088.1 GI:18642431
KEYWORDS rhl gene.
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM Plasmodium falciparum

REFERENCE 1 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
AUTHORS Taylor H.M., Grainger M. and Holder A.A.
TITLE Variation in the expression of a Plasmodium falciparum protein
JOURNAL family implicated in erythrocyte invasion
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 427)
TITLE Taylor H.M.
JOURNAL Direct Submission
SUBMITTED (04-FEB-2002) Taylor H.M., Division of Parasitology,
National Institute for Medical Research, Mill Hill, London, NW7
1AA, UNITED KINGDOM

FEATURES
source
1..427
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/isolate="PCB1"
/db_xref="taxon:5833"
/country="Colombia"
1..426
/gene="rhl"
/gene="rhl"
/function="putative role in erythrocyte invasion"
/codon_start=1
/product="rhl protein"
/protein_id="CAD23028.1"
/db_xref="GI:18642432"
/db_xref="UniProt/TREMBL:Q8T8F1"
/translation="LENIKNENIFDIOCKKKDIDIIININNTKETYLKEINRKKM
LONKQVDEKEINHHHTLOHONGVQCKKIKDHNLTTPNNSSSESHQNEQKQ
MKNILEQGTNIKPHVHNHNHNHNHNHNHNHNQXQ"

gene
CDS

ORIGIN

Alignment Scores:

Pred. No.: 0.00355 Length: 427
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-858-332G-15 (1-12) x PFA430088 (1-427)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 358 CATATCATATCATATCATATCATATCATATCATAT 393

RESULT 7

BTU49779/c 478 bp mRNA linear MAM 16-AUG-2001
LOCUS BTU49779
DEFINITION Bos taurus immunoglobulin rearranged heavy chain variable region
mRNA, partial cds.
U49779
ACCESSION U49779.1 GI:1293603
VERSION
KEYWORDS
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.

REFERENCE 1 (bases 1 to 478)
AUTHORS Berens S.J., Wylie D.E. and Lopez O.J.
TITLE Use of a single VH family and long CDR3s in the variable region of
JOURNAL Int. Immunol. 9 (1), 189-199 (1997)
PUBMED 97196871
MEDLINE 9043960
JOURNAL 2 (bases 1 to 478)
AUTHORS Berens S.J., Wylie D. and Lopez O.J.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-1996) Osvaldo J. Lopez, Biological Sciences,
University of Nebraska, 325 Mantel Hall, Lincoln, NE 68588-0118,

USA

FEATURES
source

Location/Qualifiers
1..478
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/cell_type="heterohybridoma"
/tissue_type="spleen"
/dev_stage="150 day fetus"
1..478
/codon_start=2
/product="immunoglobulin heavy chain variable region"
/protein_id="AA98660.1"
/db_xref="GI:1293604"
/translation="ILGVVTLCSKRNPPMTLLPVLSAPRGVLSVOLRESGPELTP
SQTLSLTCTGSGFSLSDNSVGVRQPEKALEMIGVYAGSGTGYNAKLSRLSTSD
NSKQVSLSLSSVTTEDPATYYCARVGCISGYGCGGYGADWAGGLVTSS"

CDS

ORIGIN

Alignment Scores:

Pred. No.: 0.00399 Length: 478
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BTU49779 (1-478)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 437 CATACCATACCAACCATACCATACCATACCATAC 402

RESULT 8

AF187065 519 bp mRNA linear ROD 11-JUN-2000
LOCUS AF187065
DEFINITION Rattus norvegicus p75NTR-associated cell death executor (Nade)
mRNA, complete cds.
AF187065
ACCESSION AF187065.1 GI:8452895
VERSION
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 519)
AUTHORS Mukai J., Hachiya T., Shoji Hoshino S., Kimura M.T., Nadano D.,
Suvanto P., Hanaoka T., Li Y., Irie S., Greene L.A. and Sato T.A.
TITLE Nade, a p75NTR-associated cell death executor, is involved in
JOURNAL signal transduction mediated by the common neurotrophin receptor
p75NTR
JOURNAL J. Biol. Chem. 275 (23), 17566-17570 (2000)

REFERENCE

2 (bases 1 to 519)
AUTHORS Mukai J., Hachiya T., Hoshino S., Kimura M., Nadano D.,
Hanaoka T., Li Y., Irie S. and Sato T.A.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia
University, 630 West 168th St., P&S 11-451, New York, NY 10032, USA
FEATURES
source
1..519
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
1..519
/gene="Nade"
44..406
/gene="Nade"
/note="involved in the common neurotrophin receptor
p75NTR-mediated signal transduction; Nade"

[illegible]

Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nambavali,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL ac: <http://image.lnl.gov>
 Series: IRAC plate: 30 Row: P Column: 24
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 21314829.
 Location/Qualifiers

FEATURES

source

1. 1949
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:25862 IMAGE:4195889"
 /tissue_type="Liver, normal. 5 month old male mouse."
 /clone_id="NCI CGAP_L19"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

gene

1. 1949
 /gene="Dp111"
 /db_xref="LOCUSID:70335"
 /db_xref="MGI:1917585"
 74. 598

CDS

/gene="Dp111"
 /codon_start=1
 /product="Dp111 protein"
 /protein_id="AA029741.1"
 /db_xref="GI:20988224"
 /db_xref="LOCUSID:70335"
 /db_xref="MGI:1917585"
 /translation="MDGLRQRFERFLERQKNVATEALGALRTEGVEKRYLAAGALL
 GLYLFGYASLNCVIGFYPAVASVKAIESPKEDDTWLTWVYVAFGLVEFES
 DLLEFMPFPYAGKCAFLFCMTGPGMNGALLHYRITRPLFKHMLDLSAASQSG
 RAIDLAAIGTRDK"
 245. 505
 /gene="Dp111"
 /note="TB2 Dp1 HVA22; Region: TB2/Dp1, HVA22 family. This
 family includes members from a wide variety of eukaryotes.
 It includes the TB2/Dp1 (deleted in polyploids) protein,
 which in humans is deleted in severe forms of familial
 adenomatous polyposis, an autosomal dominant oncological
 inherited disease. The family also includes the plant
 protein of known similarity to TB2/Dp1, the HVA22 abscisic
 acid-induced protein, which is thought to be a regulatory
 protein"
 /db_xref="CDD:pfam03134"

misc_feature

ORIGIN

Alignment Scores:

Pred. No.: 0.0172 Length: 1949
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-09-858-332g-15 (1-12) x BC029741 (1-1949)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12

DB 776 CACAAACCAACCAACCAACCAACCAACCAAC 811

RESULT 17

AB039933

LOCUS AB039933 2011 bp mRNA linear ROD 14-APR-2000
 DEFINITION Mus musculus Dp111 mRNA for polyposis locus protein 1-like 1 (TB2
 protein-like 1), complete cds.

ACCESSION

AB039933

VERSION AB039933.1 GI:7573634
 KEYWORDS polyploids locus protein 1-like 1.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE

1 (bases 1 to 2011)
 Sato, H.
 Mus musculus mRNA for TB2/Dp1-like 1, complete cds
 TITLE Published only in Database (2000)
 JOURNAL 2 (bases 1 to 2011)
 REFERENCE Sato, H.
 TITLE Direct Submission
 JOURNAL Submitted (25-FEB-2000) Hajime Sato, Tohoku University School of
 Medicine, Department of Ophthalmology, Setryo-machi 1-1, Sendai,
 Miyagi 980-8574, Japan (E-mail: hassatooph.med.tohoku.ac.jp,
 Tel:81-22-717-7294, Fax:81-22-717-7298)

FEATURES

source

1. 2011
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="PSK138"
 /chromosome="10"
 /clone="PSK138"
 /tissue_type="retina"
 /dev_stage="adult"
 /note="vector:pBluescriptSK(-)"
 1. 2011
 /gene="Dp111"
 62. 667
 /gene="Dp111"
 /standard_name="deleted in polyploids 1-like 1"
 /codon_start=1
 /product="polyploids locus protein 1-like 1 (TB2
 protein-like 1)"
 /protein_id="BAA94544.1"
 /db_xref="GI:7573635"
 /translation="MDGLRQRFERFLERQKNVATEALGALRTEGVEKRYLAAGALL
 GLYLFGYASLNCVIGFYPAVASVKAIESPKEDDTWLTWVYVAFGLVEFES
 DLLEFMPFPYAGKCAFLFCMTGPGMNGALLHYRITRPLFKHMLDLSAASQSG
 RAIDLAAIGTRDVLQALARGALVTPTASTSEPPAULEDPK"
 1968. 1973
 /gene="Dp111"

gene

CDS

polysigma
 /gene="Dp111"

ORIGIN

Alignment Scores:

Pred. No.: 0.0178 Length: 2011
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-09-858-332g-15 (1-12) x AB039933 (1-2011)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12

DB 845 CACAAACCAACCAACCAACCAACCAACCAAC 880

RESULT 18

BC083830

LOCUS BC083830 2152 bp mRNA linear ROD 12-OCT-2004
 DEFINITION Rattus norvegicus cDNA clone MGC:94922 IMAGE:7112359, complete cds.
 ACCESSION BC083830
 VERSION BC083830.1 GI:54035486
 KEYWORDS MGC.

SOURCE

ORGANISM Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2152)

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Bucrow, K.H., Schaefer, C.F., Bat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, M., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stachleiron, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedlin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loguercio, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosek, S.A., McSwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wollam, D.K., Hale, S., Garcia, A.M., Gay, L.J., Halys, S.W., Fahy, J., Helton, E., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

2 (bases 1 to 2152)
Director MGC Project.
Submitted (01-OCT-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Howard Jacobs
CDNA Library Preparation: Express Genomics
DNA Sequencing By: The I.M.A.G.E. Consortium (LNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-sbnc.stanford.edu>
Contact: (Dickson, Mark) mcdpax11.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

FEATURES

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAX Plate: 182 Row: h Column: 17
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

source

Location/Qualifiers
1..2152

/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:94922 IMAGE:7112359"
/issue_type="Testis, Rat (Brown Norway)"
/clone_lib="NIH_MGC_237"
/lab_host="DH10B"
/note="Vector: pExpress1"
28..663
/codon_start=1
/product="Unknown (protein for MGC:94922)"
/protein_id="AAH83830.1"
/db_xref="GI:54035487"

CDS

translation="MDGLRQREPERLEQGNVATDAGALREARQVGRKRYAAGATLTLGLLLFGGASLLCNVIGVYPAVASVAISPNEDTWTLLVWVYALFGVAFPSDILLFWFPFYKCAFLFCMTGPNAGALLVHRVRLFLKHNVALDSASQJLSGRLADIAAGITRDVQLARGRITLVPAASSPALBPDPKSSOTLLKHK"

ORIGIN

Alignment Scores:
Pred. No.: 0.0191 Length: 2152
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-858-332g-15 (1-12) x BC083830 (1-2152)

Qy 1 HisAenHisAenHisAenHisAenHisAenHisAen 12
Db 919 CACATCACAATCACAATCACAATCACAATCACAAT 954

RESULT 19
AX461526/c 2270 bp DNA Linear PAT 08-JUL-2002

LOCUS AX461526
DEFINITION Sequence 455 from Parent WO0198480.
ACCESSION AX461526
VERSION AX461526.1 GI:21726737
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopses.

REFERENCE 1
Budworth, P., Brown, D., Chang, H.S., Zhu, T., Han, B., Wang, X. and Cooper, B.
Promoters for regulation of plant gene expression
Patent: WO 0198480-A 455 27-DEC-2001;
Syngenta Participations AG (CH)

TITLE Location/Qualifiers
JOURNAL 1..2270
source /organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 0.0202 Length: 2270
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-858-332g-15 (1-12) x AX461526 (1-2270)

Qy 1 HisAenHisAenHisAenHisAenHisAenHisAen 12
Db 604 CACATCACAATCACAATCACAATCACAATCACAAT 569

RESULT 20
AF020407 2396 bp DNA linear INV 11-MAY-1998
LOCUS AF020407
DEFINITION Dictyostelium discoideum wima (wima) gene, partial cds.
ACCESSION AF020407 AF012944
VERSION AF020407.1 GI:2425142
KEYWORDS
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE 1 (bases 1 to 2396)
AUTHORS Loomis, W.F. and Iranfar, N.
TITLE Direct Substitution
JOURNAL Submitted (20-AUG-1997) Biology, UCSD, 9500 Gilman Drive, La Jolla, CA 92093, USA
COMMENT On May 11, 1998 this sequence version replaced gi:2393950.
FEATURES
source Location/Qualifiers
1..2396
/organism="Dictyostelium discoideum"
/mol_type="genomic DNA"
/strain="AX4"
/db_xref="taxon:44689"
702..2396
/gene="wima"

Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L21360
Center clone name: 195_K_9

* NOTE: This record contains 81 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
675 674: contig of 674 bp in length
775 774: gap of 100 bp
1481 1480: contig of 706 bp in length
1581 1580: gap of 100 bp
2302: contig of 722 bp in length
2402: gap of 100 bp
2403 3131: contig of 729 bp in length
3132 3231: gap of 100 bp
3232 3872: contig of 641 bp in length
3972: gap of 100 bp
3973 4671: contig of 699 bp in length
4672: gap of 100 bp
4771 5477: contig of 706 bp in length
4772 5577: gap of 100 bp
5478 6284: contig of 707 bp in length
5578 6384: gap of 100 bp
6285 7076: contig of 692 bp in length
7077 7176: gap of 100 bp
7177 7885: contig of 709 bp in length
7886 7985: gap of 100 bp
7986 8672: contig of 687 bp in length
8673 8772: gap of 100 bp
8773 9474: contig of 702 bp in length
9475 9574: gap of 100 bp
9575 10310: contig of 736 bp in length
10311 10410: gap of 100 bp
10411 11129: contig of 719 bp in length
11130 11229: gap of 100 bp
11229 11939: contig of 710 bp in length
11939 12039: gap of 100 bp
12040 12747: contig of 708 bp in length
12748 12847: gap of 100 bp
12848 13556: contig of 709 bp in length
13557 13656: gap of 100 bp
13657 14365: contig of 709 bp in length
14366 14465: gap of 100 bp
14466 15179: contig of 714 bp in length
15180 15279: gap of 100 bp
15280 15988: contig of 709 bp in length
15989 16088: gap of 100 bp
16089 16763: contig of 675 bp in length
16764 16863: gap of 100 bp
16864 17576: contig of 713 bp in length
17577 17676: gap of 100 bp
17677 18391: contig of 715 bp in length
18392 18491: gap of 100 bp
18492 19051: contig of 560 bp in length
19052 19151: gap of 100 bp
19152 19864: contig of 713 bp in length
19865 19964: gap of 100 bp
19965 20675: contig of 711 bp in length
20676 20775: gap of 100 bp
20776 21488: contig of 713 bp in length
21489 22307: contig of 719 bp in length
22308 22407: gap of 100 bp

22408 23129: contig of 722 bp in length
23130 23229: gap of 100 bp
23230 23941: contig of 712 bp in length
23942 24041: gap of 100 bp
24042 24729: contig of 688 bp in length
24730 24829: gap of 100 bp
24830 25548: contig of 719 bp in length
25549 25649: gap of 100 bp
25649 26389: contig of 740 bp in length
26389 26489: gap of 100 bp
26489 27215: contig of 727 bp in length
27216 27315: gap of 100 bp
27316 28022: contig of 707 bp in length
28023 28122: gap of 100 bp
28123 28836: contig of 714 bp in length
28837 28936: gap of 100 bp
28937 29646: contig of 710 bp in length
29647 29746: gap of 100 bp
29747 30455: contig of 709 bp in length
30456 30555: gap of 100 bp
30556 31269: contig of 714 bp in length
31270 31369: gap of 100 bp
31370 32086: contig of 717 bp in length
32087 32186: gap of 100 bp
32187 32905: contig of 719 bp in length
32906 33005: gap of 100 bp
33006 33722: contig of 717 bp in length
33723 33822: gap of 100 bp
33823 34502: contig of 680 bp in length
34503 34602: gap of 100 bp
34602 35297: contig of 695 bp in length
35298 35397: gap of 100 bp
35398 36118: contig of 721 bp in length
36119 36218: gap of 100 bp
36219 36930: contig of 712 bp in length
37030: gap of 100 bp
37031 37730: contig of 700 bp in length
37731 37830: gap of 100 bp
37831 38533: contig of 703 bp in length
38534 38633: gap of 100 bp
38633 39346: contig of 713 bp in length
39347 39446: gap of 100 bp
39447 40156: contig of 710 bp in length
40157 40256: gap of 100 bp
40257 40954: contig of 698 bp in length
40955 41054: gap of 100 bp
41054 41765: contig of 711 bp in length
41766 41865: gap of 100 bp
41866 42599: contig of 734 bp in length
42600 42699: gap of 100 bp
42700 43424: contig of 725 bp in length
43425 43524: gap of 100 bp
43525 44232: contig of 708 bp in length
44233 44332: gap of 100 bp
44333 45042: contig of 710 bp in length
45043 45142: gap of 100 bp
45143 45846: contig of 704 bp in length
45847 45946: gap of 100 bp
45947 46654: contig of 708 bp in length
46655 46754: gap of 100 bp
46755 47467: contig of 713 bp in length
47468 47567: gap of 100 bp
47568 48283: contig of 716 bp in length
48284 48383: gap of 100 bp
48384 49091: contig of 708 bp in length
49092 49191: gap of 100 bp
49192 49895: contig of 704 bp in length
49896 49995: gap of 100 bp
49996 50672: contig of 677 bp in length
50673 50772: gap of 100 bp
50773 51486: contig of 714 bp in length
51487 51586: gap of 100 bp
51587 52304: contig of 718 bp in length

5289 5388: gap of 100 bp
5389 5389: contig of 788 bp in length
6177 6276: gap of 100 bp
6277 7069: contig of 793 bp in length
7070 7169: gap of 100 bp
7170 7962: contig of 793 bp in length
7963 8062: gap of 100 bp
8063 8850: contig of 788 bp in length
8851 8950: gap of 100 bp
8951 9748: contig of 798 bp in length
9749 9848: gap of 100 bp
9849 10635: contig of 787 bp in length
10636 10735: gap of 100 bp
10736 11533: contig of 798 bp in length
11534 11633: gap of 100 bp
11634 12431: contig of 798 bp in length
12432 12531: gap of 100 bp
12532 13307: contig of 776 bp in length
13308 13407: gap of 100 bp
13408 14170: contig of 763 bp in length
14171 14270: gap of 100 bp
14271 15067: contig of 797 bp in length
15068 15167: gap of 100 bp
15168 15917: contig of 750 bp in length
15918 16017: gap of 100 bp
16018 16837: contig of 820 bp in length
16838 16937: gap of 100 bp
16939 17739: contig of 802 bp in length
17740 18651: contig of 812 bp in length
18652 18751: gap of 100 bp
18752 19541: contig of 790 bp in length
19542 19641: gap of 100 bp
19642 20440: contig of 799 bp in length
20441 20540: gap of 100 bp
20541 21351: contig of 811 bp in length
21352 21451: gap of 100 bp
21452 22245: contig of 794 bp in length
22246 22345: gap of 100 bp
22346 23146: contig of 801 bp in length
23147 23246: gap of 100 bp
23247 24029: contig of 783 bp in length
24030 24129: gap of 100 bp
24130 24881: contig of 752 bp in length
24882 24981: gap of 100 bp
24982 25777: contig of 796 bp in length
25778 25877: gap of 100 bp
25878 26670: contig of 793 bp in length
26671 26770: gap of 100 bp
26771 27591: contig of 821 bp in length
27592 27691: gap of 100 bp
27692 28497: contig of 806 bp in length
28498 28597: gap of 100 bp
28598 29408: contig of 811 bp in length
29409 29508: gap of 100 bp
29509 30291: contig of 783 bp in length
30292 30391: gap of 100 bp
30392 31198: contig of 807 bp in length
31199 31298: gap of 100 bp
31299 32094: contig of 796 bp in length
32095 32194: gap of 100 bp
32195 32995: contig of 801 bp in length
32996 33095: gap of 100 bp
33096 33898: contig of 803 bp in length
33899 34797: contig of 799 bp in length
34798 34897: gap of 100 bp
34898 35672: contig of 775 bp in length
35673 35772: gap of 100 bp
35773 36581: contig of 809 bp in length
36582 37495: contig of 814 bp in length
37496 37595: gap of 100 bp

37596 38431: contig of 836 bp in length
38432 38531: gap of 100 bp
38532 39344: contig of 813 bp in length
39345 39444: gap of 100 bp
39445 40252: contig of 818 bp in length
40253 40363: gap of 100 bp
40363 41168: contig of 806 bp in length
41169 41268: gap of 100 bp
41269 42068: contig of 800 bp in length
42069 42169: gap of 100 bp
42169 42960: contig of 792 bp in length
42961 43060: gap of 100 bp
43061 43853: contig of 793 bp in length
43854 43953: gap of 100 bp
43953 44719: contig of 766 bp in length
44720 44819: gap of 100 bp
44819 45625: contig of 806 bp in length
45625 45725: gap of 100 bp
45725 46549: contig of 824 bp in length
46549 46649: gap of 100 bp
46649 47449: contig of 800 bp in length
47450 47549: gap of 100 bp
47550 48372: contig of 823 bp in length
48372 48472: gap of 100 bp
48473 49285: contig of 813 bp in length
49285 49385: gap of 100 bp
49385 50171: contig of 786 bp in length
50172 50271: gap of 100 bp
50272 51062: contig of 791 bp in length
51062 51162: gap of 100 bp
51163 51969: contig of 807 bp in length
51970 52069: gap of 100 bp
52070 52867: contig of 798 bp in length
52867 52967: gap of 100 bp
52968 53750: contig of 783 bp in length
53751 53850: gap of 100 bp
53851 54641: contig of 791 bp in length
54642 54741: gap of 100 bp
54742 55550: contig of 809 bp in length
55551 55650: gap of 100 bp
55651 56472: contig of 822 bp in length
56473 56572: gap of 100 bp
56573 57375: contig of 803 bp in length
57376 57475: gap of 100 bp
57476 58297: contig of 822 bp in length
58298 58397: gap of 100 bp
58398 59210: contig of 813 bp in length
59211 59310: gap of 100 bp
59311 60115: contig of 805 bp in length
60116 60215: gap of 100 bp
60216 61021: contig of 806 bp in length
61022 61121: gap of 100 bp
61122 61907: contig of 786 bp in length
61908 62007: gap of 100 bp
62008 62797: contig of 790 bp in length

Alignment Scores:

Pred. No.: 0.816 Length: 79776
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservatvie: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-858-332g-15 (1-12) x AC139431 (1-79776)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 63185 CATATCATATCATATCATATCATATCATAT 63150

RESULT 34

ATT805

ATT805

83029 bp

DNA

linear

PLN 16-FEB-1998

DEFINITION	Arabidopsis thaliana DNA chromosome 4, BAC clone T805 (ESSAII project).
ACCESSION	AL021890
VERSION	AL021890.1 GI:2894557
KEYWORDS	Arabidopsis thaliana (thale cress)
SOURCE	Arabidopsis thaliana
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE	1 Bevan, M., Wedler, H., Wambutt, R., Bancroft, I., Mewes, H.W., Mayer, K. and Scheller, C.
AUTHORS	2 (bases 1 to 83029)
JOURNAL	Unpublished
REFERENCE	EU Arabidopsis sequencing project.
AUTHORS	Direct Submission
TITLE	Submitted (13-FEB-1998) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@berc.ac.uk
JOURNAL	Location/Qualifiers
FEATURES	1. 83029
source	/organism="Arabidopsis thaliana"
	/mol_type="genomic DNA"
	/variety="Columbia"
	/db_xref="taxon:3702"
	/chromosome="4"
gene	1. 1440
	/gene="T805.10"
CDS	join(1. 79,179. 459,515. 805,893. 1098,1176. 1440)
	/gene="T805.10"
	/note="Protein sequence is in conflict with the conceptual translation, similarity to Hypothetical protein, Saccharomyces cerevisiae, MNOS:S71593"
	/codon_start=1
	/product="putative protein"
	/protein_id="CAI17147.1"
	/db_xref="GI:2894558"
	/db_xref="GAI:049703"
	/db_xref="UniProt/TREMBL:O49703"
	/translation="KLRVSAASSSSNFKKKPILIIIVGNAGSGKTSFLRLVCHTPTS KSHGVNLDPAVMSLPGANIDIRTVKKEVKOYNLGNGLISNLPAFKTDE VNIPLFLVOMITDFSGHCVSVIEKRADOLDYLVDPGQIEIFETMSAGAIITTEFA STPEPTVTVVDPRSSRPTTMSNNLYACSIIXKRLPVLAFNKTVDADHGFALFV MEDFEVFOAIQSDNSYTATLANSLSLSTFPRNIRSVGVSATSGAGMDGFKALFA SAEEMETRYKADLDNRKADKERLEERKHEMKLAKDMESSQGTVVINTGLKORDA TERKMLEDEDEDFQVDEEDSDDAIDEDDEDETKHYLL"
exon	1. 79
	/gene="T805.10"
	/number=1
intron	80. 178
	/gene="T805.10"
	/number=1
exon	179. 459
	/gene="T805.10"
	/number=2
intron	460. 514
	/gene="T805.10"
	/number=2
exon	515. 805
	/gene="T805.10"
	/number=3
intron	806. 892
	/gene="T805.10"
	/number=3
exon	893. 1098
	/gene="T805.10"
	/number=4
intron	1099. 1175
	/gene="T805.10"
	/number=4
exon	1176. 1440
	/gene="T805.10"
	/number=5
gene	1809. 4079
	/gene="T805.20"
	/complement(join(1809. 1991,2084. 2137,2228. 2311, 2409. 2623,3028. 3127,4038. 4079))
	/gene="T805.20"
CDS	complement(join(1809. 1991,2084. 2137,2228. 2311, 2409. 2623,3028. 3127,4038. 4079))
	/gene="T805.20"
	/note="similarity to NADH dehydrogenase chain 4, Acanthamoeba castellanii mitochondrion, PIR2:S53834 contains EST gb:N38447, T21918"
	/codon_start=1
	/product="putative protein"
	/protein_id="CAI17148.1"
	/db_xref="GI:2894559"
	/db_xref="UniProt/TREMBL:O49704"
	/translation="MTPDVAAAPVKRTAIIISPVNLYLNPTLVVKOYQFWRLYTNPLVF RMDIDFLHMFPLARCYKLENSFRGKTPDVLVLTGATVLTGIVLGGIIPYLS VSPSKTIFLNSLITPMYVYVSKONPYITMSFLGLFTPAAYLPMVTLGFSITLVGASA WDFPLGMLIGHAIVFLAFVYPRMTDRRLPKTSPFLKALFADEPVVIARPDVRFALAP FDEIHQD"
exon	complement(1809. 1991)
	/gene="T805.20"
	/number=1
intron	complement(1992. 2083)
	/gene="T805.20"
	/number=1
exon	complement(2084. 2137)
	/gene="T805.20"
	/number=2
intron	complement(2138. 2227)
	/gene="T805.20"
	/number=2
exon	complement(2228. 2311)
	/gene="T805.20"
	/number=3
intron	complement(2312. 2408)
	/gene="T805.20"
	/number=3
exon	complement(2409. 2623)
	/gene="T805.20"
	/number=4
intron	complement(2624. 3027)
	/gene="T805.20"
	/number=4
exon	complement(3028. 3127)
	/gene="T805.20"
	/number=5
intron	complement(3128. 4037)
	/gene="T805.20"
	/number=5
gene	4022. 10335
	/gene="T805.30"
CDS	join(4022. 4729,4822. 5290,5735. 5961,6060. 6125. 6311. 6364,6732. 6848,7056. 7154,7211. 7391,7599. 7736, 7938. 8132,8214. 8494,9384. 9606,9673. 9937,9981. 10136, 10248. 10335)
	/gene="T805.30"
	/note="similarity to microtubule associated protein, Drosophila melanogaster, PATCHX.G130122"
	/codon_start=1
	/product="putative protein"
	/protein_id="CAI17149.1"
	/db_xref="GI:2894560"
	/db_xref="UniProt/TREMBL:O49705"
	/translation="MDENEPPCASAPAPRNPASSLISDINSFKTPRRTSVNSNISK SYPHPPTAKONPKSSSSNPRPSPVHVSASKYSTSRRLKAFELQOOSRRKEL TKEKNIKSLASLTYWLNFLFENPENCDDPPEBNSSGVNGLKGRDSEBALGNSKSV GVDITMRSFKRLNIGMCGEKRSIIDSLSKYSTLRESLRDVCSDLDDLRKQWFL ISLGCKELFDVWTRVSKNIDGRIRIMKQCPVLVDPQMKERAIKALMSYNQVWRLG

LYIFGDSFLSDSEVNSDOEMAFLMWITSKOPFSDHGLARSTYINKVEGLYRPGY
EALGNVILKRIILLVLVIDRAKOSCLSLRTGIDGSDSPMFESEKSIKSHOLIC
GSRFDLMRYLKKMVTVPSPNKRKLANCRIALQYIKDGVSLKDEGMWITGEVAD
GPRELITSLIMNIFVHLQPLINGRLTTEEYIKQVGNQNNQITMSPLEMLNMIQ
DPGQGPQSVKSNSTYHDVONFIIISOQLTALGSPPEIGDLRHNVAVNSQSVITL
LAPLSKRLVKNLIDFTLAGYLTNSVFQEOQLNPFKLICSCVOYOEKRSRISCS
SSSAVTIEEPRENGEDATKRFQAIKAMQOMANOQISVGNASHITOGSLSRCTT
DSORAVILLAEIAAVIIGSNIRGLHARRKFRKMAICGLQAAVITLSTKRLQVEK
FNVEEYVTLSESRANLKPARYVYKTYDPSFPIKRSVSYIQKAVRHSNHL
KALRIQLAMRSYKERVISITIOSVYRMITRMNRTYKFSIILIQVSPYPIISQ
SHMRGYLTAKSAQYLDLRVMQITSANIDKRLINKLSALSELISMKNVNIILH
ICETLESATKISDCKCEELVAGALIKLTLIRASRSIPDOQVSKALSTGLHLY
PQWDELINTKSIQITFWEILRSSFSQVFLSPKKEBAVFIASDVLKTKISHRG
VEAVRKLPLVLRHLVVELTRKANIERKNVQGTGKESBRRLKEAIELVKLITTR

exon 4022..4729
/gene="T805.30"
/number=1
exon complement(4038..4079)
/gene="T805.20"
/number=6
intron 4730..4821
/gene="T805.30"
/number=1
exon 4822..5290
/gene="T805.30"
/number=2
intron 5291..5734
/gene="T805.30"
/number=2
exon 5735..5961
/gene="T805.30"
/number=3
intron 5962..6059
/gene="T805.30"
/number=3
exon 6060..6125
/gene="T805.30"
/number=4
intron 6126..6310
/gene="T805.30"
/number=4
exon 6311..6364
/gene="T805.30"
/number=5
intron 6365..6731
/gene="T805.30"
/number=5
exon 6732..6848
/gene="T805.30"
/number=6
intron 6849..7055
/gene="T805.30"

Alignment Scores:

Pred. No.: 0.851 Length: 83029
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-09-858-332G-15 (1-12) x ATT805 (1-83029)

QY 1 H1sAenh1sAenh1sAenh1sAenh1sAenh1sAen 12
DB 3476 CACAATCACAATCACAATCACAATCACAAT 3511

RESULT 35
CNS07T1W/C 101256 bp DNA linear PLN 24-OCT-2001
LOCUS
DEFINITION DNA centromeric region sequence from BAC DP35C12 of chromosome 5 of
Podospora anserina.

ACCESSION AL627322
VERSION AL627322.1 GI:16444879
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Submitted (24-OCT-2001) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 Evry cedex, FRANCE. [E-mail :
seq@genoscope.cns.fr - Web : www.genoscope.cns.fr]
* Christian Barreau and Beatrice Turcq, Institut de Biochimie et de
Genetique Cellulaire, UPR CNRS 9026, 1 rue Camille Saint Saens,
F-33077 Bordeaux, France;
* Alain Billault, Molecular Engines Laboratories, 20 rue Bouvier,
F-75011 Paris, France;
* Laurence Cattoir and Simone Duprat, Genoscope, Centre National
de Sequencage, 2 rue Gaston Cremieux, CP5706, F-91057 Evry CEDEX,
France; * Robert Debuchy, Sebastien Kicka and Philippe Silar,
Institut de Genetique et Microbiologie, UMR CNRS 8621, Batiment
400, Universite Paris-Sud, F-91405 Orsay CEDEX, France;
* Annie Sainsard-Chanet and Carole H. Sellem, Centre de Genetique
Moleculaire, UPR CNRS 9061, Batiment 24, allée de la terrasse,
F-91190 Gif sur Yvette, France;
base calling/assembly program : Phred/Phrap;
consensus quality: 98,93% of bases at least Q40.
FEATURES
source
1..101256
/organism="Podospora anserina"
/mol_type="genomic DNA"
/db_xref="taxon:5145"
/chromosome="5"
/clone="DP35C12"

ORIGIN

Alignment Scores:
Pred. No.: 1.05 Length: 101256
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-09-858-332G-15 (1-12) x CNS07T1W (1-101256)

QY 1 H1sAenh1sAenh1sAenh1sAenh1sAenh1sAen 12
DB 49868 CACAACCAACAACAACAACAACAATCACAATCACAAC 49833

RESULT 36
AP006378 101752 bp DNA linear PLN 22-JUL-2003
LOCUS
DEFINITION Lotus corniculatus var. japonicus genomic DNA, chromosome 1,
AP006378
AP006378.1 GI:31581009
HTG.
SOURCE
ORGANISM

Lotus corniculatus var. japonicus (Lotus japonicus)
Lotus corniculatus var. japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteeae;
Lotus.
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL

Direct Submission
Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research, 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba 292-0818, Japan (E-mail: sseto@kazusa.or.jp,
URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex. 2337),
Fax: 81-438-52-3934)

FEATURES

Location/Qualifiers
1. 101752
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"
/chromosome="1"
/clone="LjT44A10"
/clone_lib="LjT library"
/note="TAC clone: TM0223-synonym: Lotus japonicus"

ORIGIN

Alignment Scores:

Pred. No.: 1.05 Length: 101752
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-09-858-332g-15 (1-12) x AP006378 (1-101752)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 33052 CACATCATCATCATCATCATCATCATCATCAT 33087

RESULT 37

AC109729_3/c

WPCOMMENT

Sequence split into 5 fragments LOCUS AC109729 Accession AC109729

Fragment Name

Fragment Name	Begin	End
AC109729_0	1	110000
AC109729_1	100001	210000
AC109729_2	200001	310000
AC109729_3	300001	410000
AC109729_4	400001	427537

Continuation (4 of 5) of AC109729 from base 300001 (AC109729 Rattus norvegicus clone CH2)

Alignment Scores:

Pred. No.: 1.14 Length: 110000
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-858-332g-15 (1-12) x AC109729_3 (1-110000)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 60238 CACACCAACACACATCATCATCATCATCAT 60203

RESULT 38

AC120234_1/c

WPCOMMENT

Sequence split into 5 fragments LOCUS AC120234 Accession AC120234

Fragment Name

Fragment Name	Begin	End
AC120234_0	1	110000
AC120234_1	100001	210000
AC120234_2	200001	310000
AC120234_3	300001	410000
AC120234_4	400001	444501

Continuation (2 of 5) of AC120234 from base 100001 (AC120234 Rattus norvegicus clone CH2)

Alignment Scores:

Pred. No.: 1.14 Length: 110000
Score: 84.00 Matches: 12

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-858-332g-15 (1-12) x AC120234_1 (1-110000)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 91217 CACACCAACACCAACACCAACACCAACAC 91182

RESULT 39

AC141230_3/c

WPCOMMENT

Sequence split into 7 fragments LOCUS AC141230 Accession AC141230

Fragment Name

Fragment Name	Begin	End
AC141230_0	1	110000
AC141230_1	100001	210000
AC141230_2	200001	310000
AC141230_3	300001	410000
AC141230_4	400001	510000
AC141230_5	500001	610000
AC141230_6	600001	630326

Continuation (4 of 7) of AC141230 from base 300001 (AC141230 Homo sapiens chromosome 16)

Alignment Scores:

Pred. No.: 1.14 Length: 110000
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-858-332g-15 (1-12) x AC141230_3 (1-110000)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 99966 CATATCATATCATATCATATCATATCATAT 99931

RESULT 40

CR382132_16/c

WPCOMMENT

Sequence split into 40 fragments LOCUS CR382132 Accession CR382132

Fragment Name

Fragment Name	Begin	End
CR382132_00	1	110000
CR382132_01	100001	210000
CR382132_02	200001	310000
CR382132_03	300001	410000
CR382132_04	400001	510000
CR382132_05	500001	610000
CR382132_06	600001	710000
CR382132_07	700001	810000
CR382132_08	800001	910000
CR382132_09	900001	1010000
CR382132_10	1000001	1110000
CR382132_11	1100001	1210000
CR382132_12	1200001	1310000
CR382132_13	1300001	1410000
CR382132_14	1400001	1510000
CR382132_15	1500001	1610000
CR382132_16	1600001	1710000
CR382132_17	1700001	1810000
CR382132_18	1800001	1910000
CR382132_19	1900001	2010000
CR382132_20	2000001	2110000
CR382132_21	2100001	2210000
CR382132_22	2200001	2310000
CR382132_23	2300001	2410000
CR382132_24	2400001	2510000
CR382132_25	2500001	2610000
CR382132_26	2600001	2710000
CR382132_27	2700001	2810000
CR382132_28	2800001	2910000

CR382132_29 2900001 3010000
 CR382132_30 3000001 3110000
 CR382132_31 3100001 3210000
 CR382132_32 3200001 3310000
 CR382132_33 3300001 3410000
 CR382132_34 3400001 3510000
 CR382132_35 3500001 3610000
 CR382132_36 3600001 3710000
 CR382132_37 3700001 3810000
 CR382132_38 3800001 3910000
 CR382132_39 3900001 4003362
 Continuation (17 of 40) of CR382132 from base 1600001 (CR382132 Yarrowia lipolytica chrc

Alignment Scores:

Pred. No.: 1.14 Length: 110000
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 8 Gaps: 0

US-09-858-332G-15 (1-12) x CR382132_16 (1-110000)

Qy 1 HisAenHisaEnHisaEnHisaEnHisaEn 12
 Db 25949 CACAACCAACCAACCAACCAACCAACCAAC 25914

RESULT 41
 AE016820_05/c

Sequence split into 15 fragments LOCUS AE016820 Accession AE016820

Fragment Name	Begin	End
AE016820_00	1	110000
AE016820_01	100001	210000
AE016820_02	200001	310000
AE016820_03	300001	410000
AE016820_04	400001	510000
AE016820_05	500001	610000
AE016820_06	600001	710000
AE016820_07	700001	810000
AE016820_08	800001	910000
AE016820_09	900001	1010000
AE016820_10	1000001	1110000
AE016820_11	1100001	1210000
AE016820_12	1200001	1310000
AE016820_13	1300001	1410000
AE016820_14	1400001	1476513

Continuation (6 of 15) of AE016820 from base 500001 (AE016820 Eremothecium gossypii ATCC

Alignment Scores:

Pred. No.: 1.14 Length: 110000
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 8 Gaps: 0

US-09-858-332G-15 (1-12) x AE016820_05 (1-110000)

Qy 1 HisAenHisaEnHisaEnHisaEnHisaEn 12
 Db 91897 CACAACCAACCAACCAACCAACCAACCAAC 91862

RESULT 42
 AP004065/c

LOCUS AP004065 112057 bp DNA linear HTG 21-MAR-2002
 DEFINITION Oryza sativa (japonica cultivar-group) chromosome 8 clone
 ACCESSION AP004065
 VERSION AP004065.1 GI:15208433
 KEYWORDS HTG; HTGS PHASE2
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)

REFERENCE

AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Submitted (15-AUG-2001) Takuji Sasaki, National Institute of
 Agricultural Resources, Rice Genome Research Program, Kamondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/
 Tel:81-298-38-7441, Fax:81-298-38-7468)
 The nucleotide sequence of this BAC clone was generated by
 combining Monanto and RGP-Japan sequencing data.
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 Location/Qualifiers
 1..112057

FEATURES

source
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="8"
 /clone="OJ1521_B01"

ORIGIN

Alignment Scores:
 Pred. No.: 1.16 Length: 112057
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-858-332G-15 (1-12) x AP004065 (1-112057)

Qy 1 HisAenHisaEnHisaEnHisaEnHisaEn 12
 Db 108177 CACAACCAACCAACCAACCAACCAACCAAC 108142

RESULT 43

LOCUS AC115599 125623 bp DNA linear INV 12-MAR-2003
 DEFINITION Dictyostelium discoideum chromosome 2 map 4229098-4354721 strain
 ACCESSION AC115599
 VERSION AC115599.2 GI:28828893
 KEYWORDS HTG.
 SOURCE Dictyostelium discoideum
 ORGANISM Dictyostelium discoideum

REFERENCE

AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Submitted (15-AUG-2001) Takuji Sasaki, National Institute of
 Agricultural Resources, Rice Genome Research Program, Kamondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/
 Tel:81-298-38-7441, Fax:81-298-38-7468)
 The nucleotide sequence of this BAC clone was generated by
 combining Monanto and RGP-Japan sequencing data.
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 Location/Qualifiers
 1..112057

[illegible]

COMMENT

Tel:81-298-38-7441, Fax:81-298-38-7468)

On Dec 2, 2003 this sequence version replaced gi:33468826.
Genes were predicted from the integrated results of the following:

GENSCAN (<http://CCR-081.mt.edu/GENSCAN.html>), FGENESH
(<http://www.softberry.com/>), GeneMark.hmm
(<http://opal.biology.gatech.edu/GeneMark/>), GlimmerM
(http://www.tigr.org/tdb/glimmer/glmr_form.html), RiceHMM
(<http://rgp.dna.affrc.go.jp/RiceHMM/>), SplicePredictor
(<http://bioinformatics.iastate.edu/cgi-bin/sp.cgi>), sim4
(<http://globin.cse.psu.edu/html/doces/sim4.html>), gap2
(<http://www.tigr.org/software/glimmer/>), BLASTN and BLASTX. The
genomic sequence was searched against NCBI NonRedundant Protein
database, nr (<ftp://ncbi.nlm.nih.gov/blast/db>) and the cDNA
regions were searched against NCBI NonRedundant Protein database
with BLASTX. ESTs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no. and RGP clone ID.
Full-length cDNAs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with full-length cDNA or
sequence) is classified as an 'unknown' protein. A gene predicted
by two or more gene prediction programs is classified as a
'hypothetical' protein according to IRGSP standard. A gene
predicted by a single gene prediction program is also classified as a
miscellaneous feature of the sequence.
The orientation of the sequence is from -2M13 to M13rev of the BAC
clone. This sequence of B1147B12 clone has an overlap with P0015C07
(DBJ: AP004654) clone at 3' end. Detailed information on overlap
and assembly quality together with annotation of this entry is
available at <http://rgp.dna.affrc.go.jp/Genomeseq.html>.

FEATURES

Source

1. 127902

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"

/culivar="Nipponbare"
/db_xref="taxon:39947"

/chromosome="8"
/clone="B1147B12"

2332..2511
/gene="B1147B12.1"

2332..2511
/gene="B1147B12.1"

/note="hypothetical ORF
predicted by GlimmerM
this category is not included in IRGSP standard"

3839..7735
/gene="B1147B12.2"

join(<3839..3868,5548..6742,7643..>7735)
/gene="B1147B12.2"

/note="start and end point are not identified"

join(3839..3868,6548..6742,7643..7735)
/gene="B1147B12.2"

/note="predicted by FGENESH etc."

/codon_start=1
/product="hypothetical protein"

/protein_id="BAD03544.1"
/db_xref="GI:38637283"

/translation="METDAGAMELVLPVIGPMIANSIKLFLAQCVODIHVIGIEL
ETAMNHNHOVIGDSKATINDAPNVEARSLSDKRAQCLYESVDEKSVLSRMQICV
GVN"

complement(join(8627..8640,9382..9529))
/gene="B1147B12.3"

complement(join(8627..8640,9382..9529))
/gene="B1147B12.3"

/note="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"

mRNA

/gene="B1147B12.4"

complement(join(<14133..14199,14473..14590,14671..14727,
14943..15010,15104..15144,15405..15461,15980..16041,
16633..16929,17016..17102,17171..17263,17394..17452,
17653..17716))

/gene="B1147B12.4"

/note="start and end point are not identified"

complement(join(14133..14199,14473..14590,14671..14727,
14943..15010,15104..15144,15405..15461,15980..16041,
16633..16929,17016..17102,17171..17263,17394..17452,
17653..17716))

/gene="B1147B12.4"

/codon_start=1
/product="putative zinc transporter zupT"

/protein_id="BAD03545.1"

/db_xref="GI:38637282"

/translation="MDPKVVALSLILGSLTGLALILNCAPNKXITGLMLOGFA
TGLMISWSPFLADAVNAIGFLKGNLMFEGALLFTSTIAVPEPCNLADENDKHL
EKRTGNNNIAGKELMMRRRRVSVVTAIVAGVSLONPVGTAFTGTRVGL
NLSLAIHVIPEGISVALPAYFATCSKMOAFKLATISGRAPLGVITVYLPSSIN
PEILLEGVLGVGVAAFLTYEMLPPIFEVAGRDAVAVFGAMFMSMRHERRMDT
YGLVPHHLAAQQ"

18005..19493
/gene="B1147B12.5"

join(<18005..18732,18835..19026,19132..>19493)
/gene="B1147B12.5"

/note="supported by full-length cDNA(s) : AK108701"

18005..19493
/gene="B1147B12.5"

/note="contains full-length cDNA(s) : AK108701
non-coding transcript
probably inactive due to long 5' UTR in CDS"

complement(19926..27167)
/gene="B1147B12.6"

complement(19926..27167)
/gene="B1147B12.6"

/note="contains full-length cDNA(s) : AK121153,AK102424
non-coding transcript
probably inactive due to long 3'5'UTR in CDS"

join(29795..30195,30289..30365,31668..33560,33643..34147,
34250..34358,34436..34548,34619..34806,35481..36129,
36214..36691,36760..36855,37042..37141,37301..37360,
38970..39282)

/gene="B1147B12.7-1"

join(29795..30195,30289..30365,31668..33560,33643..34147,
34250..34358,34436..34548,34619..34806,35481..36129,
36214..36691,36760..36855,37042..37141,37301..37360,
38970..39282)

/note="supported by full-length cDNA(s) : AK120880"

join(30085..30195,30289..30365,31668..33560,33643..34147,
34250..34358,34436..34548,34619..34806,35481..36129,
36214..36691,36760..36855,37042..37141,37301..37360,
38970..39282)

/note="contains full-length cDNA(s) : AK120880"

/codon_start=1
/product="putative PMP domain protein"

/protein_id="BAD03546.1"

/db_xref="GI:38637283"

/translation="MAPAGRGARAKHTTTPQGLVLAIVKVGPPWPAKSPEDM
DMPVPRVVFVFRCTREIATIALPDQPTFNKSVWMSRKOCPKYIDSFAEA
VTEIKAYDDLPKSEITTCMLPDQSEOPTEHLVKSNNDEAPRSGMGDSPSDL
NTSGAGSTEVNDIGDSCDIDSLAVKRPKPDQPPKKPVTSTKSIINHLBD
CSATVHAERLEBPKEKEINPSEFLPTPTVOIVCALVPPKSKNAERKD
NKRVAVADISRTTPGVALDTELKRSLEKSKGFKSKIMKMSVSNESKIDTKIKM
VDSKDKLARKSSSVFSSNKKPLPDSRORLSTDMRPAKRLMDRAEITVITAK
SETSIHADNEKSDALKIEKTPAIVNSIPKIGVDWTRSGSLSLPLARISQGS
EPASGPTOLSAVPTAKKYSIKENFRVGPRLAKPRRARCFDDDEEQRTPPHKS
VKSPMGQVPTFEKQSGQGVGRISSSQIGVASMKGVAEEKPKSIGSPVHEPDY
TSPMDKRGHRLQIMGRSTTSVDSISAGNTTNLADRSQSOLMAASVYKQSG
SSSKLQTPGNLSDONPDSEKVALSKENSKAKTKSGTQITATVEGRISTAMAR
RIGMDHSEKQSDFDVKAFAEPNSPVSVMGHLIAAQARNNLIIAAQGSQDCLSA
DNVLSSTPGLPGLSPGVFHIIPASRIIPESDGMQFPDSFALTEPQGVAMKLL

REFERENCE AUTHORS TITLE JOURNAL	2 (bases 1 to 153751) Baumgart, C. Direct Submission Submitted (29-MAR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany 3 (bases 1 to 153751) Baumgart, C. Direct Submission Submitted (12-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany 5 (bases 1 to 153751) Baumgart, C. Direct Submission Submitted (21-FEB-2004) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany On or before Feb 21, 2004 this sequence version replaced gi:29850309, gi:19807752, gi:19920064, gi:19807784. CDS predictions from GenBank do not necessarily reflect true genes. Further information is available from IMB Jena, Department of Genome Analysis (http://genome.imb-jena.de/dictyostelium/) and the University Cologne, Institute for Biochemistry I (http://www.uni-koeln.de/dictyostelium/project.shtml) Funding Agency : Deutsche Forschungsgemeinschaft (DFG).
REFERENCE AUTHORS TITLE JOURNAL	2 (bases 1 to 153751) Baumgart, C. Direct Submission Submitted (29-MAR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany 3 (bases 1 to 153751) Baumgart, C. Direct Submission Submitted (12-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany 5 (bases 1 to 153751) Baumgart, C. Direct Submission Submitted (21-FEB-2004) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany On or before Feb 21, 2004 this sequence version replaced gi:29850309, gi:19807752, gi:19920064, gi:19807784. CDS predictions from GenBank do not necessarily reflect true genes. Further information is available from IMB Jena, Department of Genome Analysis (http://genome.imb-jena.de/dictyostelium/) and the University Cologne, Institute for Biochemistry I (http://www.uni-koeln.de/dictyostelium/project.shtml) Funding Agency : Deutsche Forschungsgemeinschaft (DFG).
FEATURES source	Location/Qualifiers 1..153751 /organism="Dictyostelium discoideum" /mol_type="genomic DNA" /strain="HX4" /db_xref="taxon:44689" /chromosome="2" /map="complement(581327-427576)" complement(join(<351..2554,2687..2807)) /note="GenBank exon scores (in order of location ranges) : 70.76, 0.70 - GSCU_ID dd_02254" /codon_start=1 /product="hypochemical protein" /protein_id="AA052218.1" /db_xref="GI:28829702" /translation="MNSISDHNKELFWKVFRRNKYINIEIFKQESTIEYDSFTIAS LTKSVNSNNNNNNNNKLEIFFLMEKERYMEGLDVTIISISRNFEILKILNK PYSEFLVETDNDLQKYGDELEKESITLQGLQININFLIENNFOKISIPKPT IOFLINONNDNDNDKNNNNKLIKLEIKETSIKIDELINLEIEMOKNLANTF SKENFQNSNIYWFKELYEYFQVTFIHLIYYKRYEITIEGCLSTIVRNKTENIF KLISNNNNNNNNKEIIRIITKININVEYFKNENDQILNFRNVLNNLSKEYHK LLEITGNSOIKTRGLPIEKNINILDFLLDEINNKIILKQVQIDWISNMRNNRI FTPIKTIYYKRRKKNILQFQSELEIYALKKLETNNNNNNNDNLPIKLETTIWSM VYNNPGLDSKVEYELKQNFVCKIDNCKFQKTELLELVSTNLDIWMIILINGDPTK KQDILNFMNVAHLNNTSIFYPIPTINGININPKIITHKIKPOYIINIKIKMCEQSTIL ECKDILNHLFKLIDTQKTTLNQIIDKPTINNNNNNNNNNNNNNNNNNNITKIDHS KFSHIYLNKSPRLILYIESNQIDWYLLIPQ" 351..5095 /note="GenBank exon scores (in order of location ranges) : 144.31 - GSCU_ID dd_02253" /codon_start=1 /product="similar to Arabidopsis thaliana (Mouse-ear cress). Hypochemical 59.5 kDa protein" /protein_id="AA052156.1" /db_xref="GI:28829640" /translation="MNNQYKPTLSNKAELINKMLIDNGVIRNNKLIKIVLKENNFE OTENRATTTTSRINDSTVSGIGVSLKELKVDIVAKIPKSIILSHTSSINIE KYLNNNGISLILHEASLGEKSKWYQISLPKRVDPILIMDSERKSLIKAGIAD LVNDLADILNQYADVIEISLKNHPEIFGDEIKYISINFKIANSIISRAFCVDSYH GKVVADLILFHNQIARBEVHIESNDGVCNKGSIKCKHRKVVTQHTHTNSRGT HKVAGIPSSKHIHKGNCSTTTTNEEDNOTIIEEDDEHLIKYKAGVEANKEVYN
CDS	144.31 - GSCU_ID dd_02253 /codon_start=1 /product="similar to Arabidopsis thaliana (Mouse-ear cress). Hypochemical 59.5 kDa protein" /protein_id="AA052156.1" /db_xref="GI:28829640" /translation="MNNQYKPTLSNKAELINKMLIDNGVIRNNKLIKIVLKENNFE OTENRATTTTSRINDSTVSGIGVSLKELKVDIVAKIPKSIILSHTSSINIE KYLNNNGISLILHEASLGEKSKWYQISLPKRVDPILIMDSERKSLIKAGIAD LVNDLADILNQYADVIEISLKNHPEIFGDEIKYISINFKIANSIISRAFCVDSYH GKVVADLILFHNQIARBEVHIESNDGVCNKGSIKCKHRKVVTQHTHTNSRGT HKVAGIPSSKHIHKGNCSTTTTNEEDNOTIIEEDDEHLIKYKAGVEANKEVYN
CDS	144.31 - GSCU_ID dd_02253 /codon_start=1 /product="similar to Arabidopsis thaliana (Mouse-ear cress). Hypochemical 59.5 kDa protein" /protein_id="AA052156.1" /db_xref="GI:28829640" /translation="MNNQYKPTLSNKAELINKMLIDNGVIRNNKLIKIVLKENNFE OTENRATTTTSRINDSTVSGIGVSLKELKVDIVAKIPKSIILSHTSSINIE KYLNNNGISLILHEASLGEKSKWYQISLPKRVDPILIMDSERKSLIKAGIAD LVNDLADILNQYADVIEISLKNHPEIFGDEIKYISINFKIANSIISRAFCVDSYH GKVVADLILFHNQIARBEVHIESNDGVCNKGSIKCKHRKVVTQHTHTNSRGT HKVAGIPSSKHIHKGNCSTTTTNEEDNOTIIEEDDEHLIKYKAGVEANKEVYN
CDS	144.31 - GSCU_ID dd_02253 /codon_start=1 /product="similar to Arabidopsis thaliana (Mouse-ear cress). Hypochemical 59.5 kDa protein" /protein_id="AA052156.1" /db_xref="GI:28829640" /translation="MNNQYKPTLSNKAELINKMLIDNGVIRNNKLIKIVLKENNFE OTENRATTTTSRINDSTVSGIGVSLKELKVDIVAKIPKSIILSHTSSINIE KYLNNNGISLILHEASLGEKSKWYQISLPKRVDPILIMDSERKSLIKAGIAD LVNDLADILNQYADVIEISLKNHPEIFGDEIKYISINFKIANSIISRAFCVDSYH GKVVADLILFHNQIARBEVHIESNDGVCNKGSIKCKHRKVVTQHTHTNSRGT HKVAGIPSSKHIHKGNCSTTTTNEEDNOTIIEEDDEHLIKYKAGVEANKEVYN
CDS	144.31 - GSCU_ID dd_02253 /codon_start=1 /product="similar to Arabidopsis thaliana (Mouse-ear cress). Hypochemical 59.5 kDa protein" /protein_id="AA052156.1" /db_xref="GI:28829640" /translation="MNNQYKPTLSNKAELINKMLIDNGVIRNNKLIKIVLKENNFE OTENRATTTTSRINDSTVSGIGVSLKELKVDIVAKIPKSIILSHTSSINIE KYLNNNGISLILHEASLGEKSKWYQISLPKRVDPILIMDSERKSLIKAGIAD LVNDLADILNQYADVIEISLKNHPEIFGDEIKYISINFKIANSIISRAFCVDSYH GKVVADLILFHNQIARBEVHIESNDGVCNKGSIKCKHRKVVTQHTHTNSRGT HKVAGIPSSKHIHKGNCSTTTTNEEDNOTIIEEDDEHLIKYKAGVEANKEVYN
CDS	144.31 - GSCU_ID dd_02253 /codon_start=1 /product="similar to Arabidopsis thaliana (Mouse-ear cress). Hypochemical 59.5 kDa protein" /protein_id="AA052156.1" /db_xref="GI:28829640" /translation="MNNQYKPTLSNKAELINKMLIDNGVIRNNKLIKIVLKENNFE OTENRATTTTSRINDSTVSGIGVSLKELKVDIVAKIPKSIILSHTSSINIE KYLNNNGISLILHEASLGEKSKWYQISLPKRVDPILIMDSERKSLIKAGIAD LVNDLADILNQYADVIEISLKNHPEIFGDEIKYISINFKIANSIISRAFCVDSYH GKVVADLILFHNQIARBEVHIESNDGVCNKGSIKCKHRKVVTQHTHTNSRGT HKVAGIPSSKHIHKGNCSTTTTNEEDNOTIIEEDDEHLIKYKAGVEANKEVYN
CDS	144.31 - GSCU_ID dd_02253 /codon_start=1 /product="similar to Arabidopsis thaliana (Mouse-ear cress). Hypochemical 59.5 kDa protein" /protein_id="AA052156.1" /db_xref="GI:28829640" /translation="MNNQYKPTLSNKAELINKMLIDNGVIRNNKLIKIVLKENNFE OTENRATTTTSRINDSTVSGIGVSLKELKVDIVAKIPKSIILSHTSSINIE KYLNNNGISLILHEASLGEKSKWYQISLPKRVDPILIMDSERKSLIKAGIAD LVNDLADILNQYADVIEISLKNHPEIFGDEIKYISINFKIANSIISRAFCVDSYH GKVVADLILFHNQIARBEVHIESNDGVCNKGSIKCKHRKVVTQHTHTNSRGT HKVAGIPSSKHIHKGNCSTTTTNEEDNOTIIEEDDEHLIKYKAGVEANKEVYN
CDS	144.31 - GSCU_ID dd_02253 /codon_start=1 /product="similar to Arabidopsis thaliana (Mouse-ear cress). Hypochemical 59.5 kDa protein" /protein_id="AA052156.1" /db_xref="GI:28829640" /translation="MNNQYKPTLSNKAELINKMLIDNGVIRNNKLIKIVLKENNFE OTENRATTTTSRINDSTVSGIGVSLKELKVDIVAKIPKSIILSHTSSINIE KYLNNNGISLILHEASLGEKSKWYQISLPKRVDPILIMDSERKSLIKAGIAD LVNDLADILNQYADVIEISLKNHPEIFGDEIKYISINFKIANSIISRAFCVDSYH GKVVADLILFHNQIARBEVHIESNDGVCNKGSIKCKHRKVVTQHTHTNSRGT HKVAGIPSSKHIHKGNCSTTTTNEEDNOTIIEEDDEHLIKYKAGVEANKEVYN
CDS	144.31 - GSCU_ID dd_02253 /codon_start=1 /product="similar to Arabidopsis thaliana (Mouse-ear cress). Hypochemical 59.5 kDa protein" /protein_id="AA052156.1" /db_xref="GI:28829640" /translation="MNNQYKPTLSNKAELINKMLIDNGVIRNNKLIKIVLKENNFE OTENRATTTTSRINDSTVSGIGVSLKELKVDIVAKIPKSIILSHTSSINIE KYLNNNGISLILHEASLGEKSKWYQISLPKRVDPIL


```

6.21, 14.57 - GSCJ_ID dd_02249"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA052160.1"
/db_xref="gi:28829644"
/translation="MTYDSDRYNTLRGSKQDPEFTIOEVLASAKISKQNTILINN
AEDGSVHFPSEERKAAYSKFTFLRLIDEDVLSTNVLPIDPSTDLFANCTGVLLK"
16650..17525

```

CDS	complement(join(17735. .18959,19168. .19640,19750. .20190,
-----	--

```

/note=/GeneID exon scores (in order of location ranges) :
130.29, 48.89, 51.03, 20.75 - GSCC_ID dd_00831"
/codon_start=1
/product="similar to Dictyostelium discoideum (slime
mold). Alg-2 interacting protein X"
/protein_id="AA052162.1"
/db_xref="GI:28829646"
/translation="MSSIEKKRREKVDFSKPLTKYIKQFSKASDQHTQATLNGI
REDVNPQIETLESKEMVYKYSILSELRPIISNNVRYSFPMTDYSRQRKTLTSC
YFPRFASVLEFVSGYSVQASINENIGVKKACQOPLAGVGVKTLREVASLTPC
STSDSPESSEQLMTITVAQIQECTYERKASDNDLSLSTSLAQAVAYVDPTNOLL
SKNVLATPIALEKEIVERYVSYITFYCSAKNDITVDYDTIPAHKLTPIEKKPLA

```

Alignment Scores:	
Pred. No.:	1.62
Score:	84.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	3
Length:	153751
Matches:	12
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-858-332G-15 (1-12) X AC116551 (1-153751)

Dy 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
| | | | |
Db 73027 CACAACCAACCAACCAACCAACCAAC 72992

RESULT	48
AC142155/c	
LOCUS	
DEFINITION	
AC142155	157087 bp DNA linear HTG-24-MAR-2003
Rattus norvegicus clone CH230-469L20,	WORKING DRAFT SEQUENCE, 41
unordered pieces.	

ACCESSION	AC142155
VERSION	AC142155.2
KEYWORDS	HTG, HTGS_PHASE1, HTGS_DRAFT.
SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus

REFERENCE	AUTHORS
1 (bases 1 to 157087)	Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,

Muzny, D., Meezker, M., Lee, S., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooke, S., Amin, A., Apicuno, D., Anylebsch, V., Aoyagi, A., Ayodji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bismail, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Butcher, P., Burrell, K., Calderon, B., Cardenas, V., Carter, K., Cavazos, I., Cesari, H., Centler, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chiu, J., Chiu, J., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 157087)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (23-MAR-2003)
REFERENCE	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
AUTHORS	3 (bases 1 to 157087)
TITLE	Worley, K.C.
JOURNAL	Direct Submission
REFERENCE	Submitted (24-MAR-2003)
AUTHORS	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

Delgado, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Dellado, O., Denison, S., Detamco, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durning, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabist, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregregios, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunnarute, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idolbird, R., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Loreshueta, L., Loulesged, H., Lozado, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., Meleod, M., McNeill, T., Meenen, E., Milosavljevic, A.,
Miner, G., Minis, E., Montemayor, J., Moore, S., Morgan, M., Morris, K.,
Morris, S., Mundaas, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,
Newton, N., Nguyen, N., Norris, S., Nwaoketeme, O., Okunolu, G.,
Olazunrinsagoun, A., Pal, S., Parks, K., Paternak, S., Paul, H.,
Perez, A., Perez, L., Pfankoch, C., Plopper, F., Poindexter, A.,
Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Qutroz, J., Rachlin, E.,
Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,
Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A.,
Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S.,
Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A.,
Sisson, I., Sitter, C.D., Snajs, D., Speed, A., Sodergren, E.,
Song, X.-Z., Sorelle, R., Sosa, J., Steime, M., Strong, R., Sutton, A.,
Svatek, A., Tabors, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,
Tingley, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D.,
Walden, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J.,
Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczky, R.,
Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,
Yen, J., Yoon, K., Yoon, V., Yu, P., Zhang, J., Zhou, J., Zhou, X.,
Zhao, S., Dunn, D., von Niederhauser, A., Weiss, R., Smith, D.R.,
Holt, R.A., Smith, H.O., Weinstein, G., and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 157087)
Worley, K.C.
Direct Submission
Submitted (23-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 157087)
Worley, K.C.
Direct Submission
Submitted (24-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 24, 2003 this sequence version replaced gi:29164535.

----- Genome Center -----
Center: Baylor College of Medicine
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: KELO
Center clone name: CH230-469L20
----- Summary Statistics -----
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 150122 bases at least Q40
Consensus quality: 140150 bases at least Q30
Consensus quality: 144654 bases at least Q20
Estimated insert size: 142503; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length

SWISSPROT, Tr, TREMBL, Wp, WORMPEP, Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr10> RP11-548K23 is from the library RP11-548K23 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6

FEATURES

source
1. 157193
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-548K23"
/clone_1ib="RP11-11.2"

ORIGIN

Alignment Scores:

Pred. No.: 1.65 Length: 157193
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-858-332G-15 (1-12) X AL355315 (1-157193)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 55173 CATATCATATCATATCATATCATATCATATCATAT 55208

RESULT 50

AC117110

LOCUS

DEFINITION

AC117110

AC117110

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 161701)

Muzny, D., Marie, J., Metzger, M., Lee, J., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alshrocks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenship, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, C., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Bayes, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabiella, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregergs, E., Geer, K., Gill, R., Gilly, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,

COMMENT

REFERENCE
AUTHORS
TITLE
JOURNAL

TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

TITLE
JOURNAL

-----Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 157709: contig of 157709 bp in length
 * 157710 157809: gap of unknown length
 * 157810 161701: contig of 3892 bp in length.
 Location/Qualifiers

1. 161701
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-234D18"
 complement(609..1485)
 /note="clone_boundary"
 clone_end:77
 site:

misc_feature

end_sequence:RWB0X21TU"
 complement(149221..150145)
 /note="clone_boundary"
 clone_end:Sp6
 site:

misc_feature

end_sequence:RWB0X21TV"
 156629..157709
 /note="wgs_end_extension"
 clone_end:Sp6"

misc_feature

157810..158835
 /note="wgs_end_extension"
 clone_end:Sp6"
 160542..161701
 /note="wgs_end_extension"
 clone_end:Sp6"

misc_feature

160542..161701
 /note="wgs_end_extension"
 clone_end:Sp6"

ORIGIN

Alignment Scores:
 Pred. No.: 1.7 Length: 161701
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-858-332G-15 (1-12) x AC117110 (1-161701)

OY 1 H1SAsNH1SAsNH1SAsNH1SAsNH1SAsNH1SAsN 12

Db 65903 CACACACACACACACACACACACACACACACAC 65938

RESULT 51

AC148008/c

LOCUS AC148008 161965 bp DNA linear ROD 29-MAY-2004
 DEFINITION Mus musculus BAC clone RP23-253C14 from chromosome 18, complete
 sequence.
 AC148008
 VERSION AC148008.2 GI:47131331
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 161965)
 TITLE The sequence of Mus musculus BAC clone RP23-253C14
 JOURNAL Unpublished (2001)
 REFERENCE 2 (bases 1 to 161965)

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Wilson, R.K.
 Direct Submission
 Submitted (28-JAN-2004) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 3 (bases 1 to 161965)
 Wilson, R.K.
 Direct Submission
 Submitted (12-MAY-2004) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 4 (bases 1 to 161965)
 Wilson, R.K.
 Direct Submission
 Submitted (29-MAY-2004) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 On May 12, 2004 this sequence version replaced gi:41351729.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@wustl.edu
 ----- Summary Statistics
 Center project name: M_BA0253C14

NOTICE:

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC library has been constructed by Kazuhiro Oseegawa
 and Minako Tateno in the laboratory of Pieter de Jong
 (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or
 brain genomic DNA. The clone and detailed information can be
 obtained from Research Genetics, Inc. (<http://www.reagen.com>) or
 Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is
 overlapped by AC144794 and AC120797.

FEATURES

source

1. 161965
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="18"
 /map="18"
 /clone="RP23-253C14"
 /clone_1ib="RPCT-23"
 1603..1681
 /rpt_family="MALR"
 1934..2148
 /rpt_family="MALR"
 2430..2822
 /rpt_family="MALR"
 2903..3049
 /rpt_family="B4"
 4067..4150
 /rpt_family="MALR"
 4226..4355
 /rpt_family="ERVK"

```

repeat_region      4376..4523
/rpt_family="Alu"
repeat_region      4748..4950
/rpt_family="B2"
repeat_region      5139..5350
/rpt_family="MaLR"
repeat_region      6762..6854
/rpt_family="L1"
repeat_region      6915..7458
/rpt_family="ERVK"
repeat_region      7585..7813
/rpt_family="MaLR"
repeat_region      8150..8292
/rpt_family="ERVJ"
repeat_region      9000..10058
/rpt_family="L1"
repeat_region      10059..10436
/rpt_family="RMR14"
repeat_region      10437..10563
/rpt_family="L1"
repeat_region      10654..10739
/rpt_family="B2"
repeat_region      10740..10929
/rpt_family="ERVK"
repeat_region      11337..11506
/rpt_family="B2"
repeat_region      12767..12826
/rpt_family="MaLR"
repeat_region      12870..13089
/rpt_family="ERVJ"
repeat_region      13090..13289
/rpt_family="MaLR"
repeat_region      13300..13444
/rpt_family="Alu"
repeat_region      14035..14083
/rpt_family="ERV1"
repeat_region      15986..16411
/rpt_family="MaLR"
repeat_region      17388..17450
/rpt_family="ERV1"
repeat_region      17657..17835
/rpt_family="MaLR"
repeat_region      17836..18235
/rpt_family="MaLR"
repeat_region      18236..18276
/rpt_family="MaLR"
repeat_region      18564..19053
/rpt_family="L1"
repeat_region      20485..20742
/rpt_family="L1"
repeat_region      21352..21835
/rpt_family="MaLR"
repeat_region      21813..22021
/rpt_family="MaLR"
repeat_region      22022..22221
/rpt_family="MER2_type"
repeat_region      22373..22515
/rpt_family="MER2_type"
repeat_region      22723..23139
/rpt_family="RMR19B"
repeat_region      23132..23789
/rpt_family="RMR19B"
repeat_region      26200..26479
/rpt_family="MaLR"
repeat_region      26480..27094
/rpt_family="MaLR"
repeat_region      27193..27638
/rpt_family="MaLR"
repeat_region      27636..27880
/rpt_family="MaLR"
repeat_region      27943..27995
/rpt_family="ERV1"
repeat_region      28616..28762

```

```

/rpt_family="Alu"
repeat_region      29967..30253
/rpt_family="B4"
repeat_region      32879..32988
/rpt_family="MaLR"
repeat_region      33621..35093
/rpt_family="RMR1A"
repeat_region      35367..35625
/rpt_family="MaLR"
repeat_region      36503..36910
/rpt_family="L1"
repeat_region      36911..37054
/rpt_family="Alu"
repeat_region      37055..37070
/rpt_family="L1"
repeat_region      37061..38066
/rpt_family="L1"
repeat_region      38065..39182
/rpt_family="L1"
repeat_region      39173..39733
/rpt_family="L1"
repeat_region      39734..39924
/rpt_family="B2"
repeat_region      39925..40025
/rpt_family="L1"
repeat_region      43573..43718
/rpt_family="Alu"
repeat_region      44675..45149
/rpt_family="ERV1"
repeat_region      45199..45518
/rpt_family="L1"
repeat_region      46734..46856
/rpt_family="Alu"
repeat_region      47591..47785
/rpt_family="B2"
repeat_region      47804..47997
/rpt_family="RMR17C"
repeat_region      49007..49181
/rpt_family="B2"

```

Alignment Scores:

Pred. No.:	1.7	Length:	161965
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-858-332g-15 (1-12) x AC148008 (1-161965)

Qy 1 HisAenHisAenHisAenHisAenHisAen 12

Db 141154 CACACCAACCAACCAACCAACCAACCAACCAAC 141119

AC139383/5 163087 bp DNA linear HTG 10-MAY-2003

LOCUS AC139383/5

DEFINITION Rattus norvegicus clone CH230-232M19, *** SEQUENCING IN PROGRESS

ACCESSION AC139383

VERSION AC139383.5 GI:30521079

KEYWORDS HTG; HTGS PHASRL

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 163087)

AUTHORS Muzny,D.,Marle.,Metzker,M.,Lee.,Abramson,S.,Adams,C.,Alder,J.,Allen,C.,Allen,H.,Aisbrooks,S.,Amin,A.,Anguiano,D.,Anyalebechi,V.,Aoyagi,A.,Ayodeji,M.,Baca,B.,Baden,H.,Baldwin,D.,Bandaranaike,D.,Barber,M.,Barnstead,M.,Benahmed,F.,Biswal,K.,Blair,J.,Blankenburg,K.,Blych,P.,Brown,M.,

TITLE	JOURNAL
REFERENCE	REFERENCE
AUTHORS	AUTHORS
TITLE	TITLE
JOURNAL	JOURNAL
REFERENCE	REFERENCE
AUTHORS	AUTHORS
TITLE	TITLE
JOURNAL	JOURNAL
COMMENT	COMMENT

```
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 76 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
```

1	1163:	contig of 1163 bp in length
1164	1263:	gap of unknown length
1264	2405:	contig of 1142 bp in length
2406	2505:	gap of unknown length
2506	3599:	contig of 1090 bp in length
3596	3699:	gap of unknown length
3696	4716:	contig of 1021 bp in length
4117	4816:	gap of unknown length
4817	6181:	contig of 1365 bp in length
6182	7973:	gap of unknown length
7974	8073:	contig of 1692 bp in length
8074	8638:	gap of unknown length
8639	9788:	contig of 1565 bp in length
9739	11399:	gap of unknown length
11400	11499:	contig of 1661 bp in length
11500	12507:	gap of unknown length
12508	12607:	contig of 1008 bp in length
12608	13877:	gap of unknown length
13878	13977:	contig of 1270 bp in length
13978	15121:	gap of unknown length
15122	15221:	contig of 1144 bp in length
15222	16241:	gap of unknown length
16242	16341:	contig of 1020 bp in length
16342	17364:	gap of unknown length
17365	17464:	contig of 1023 bp in length
17465	18591:	gap of unknown length
18592	18691:	contig of 1127 bp in length
18692	20093:	gap of unknown length
20094	20193:	contig of 1402 bp in length
20194	21253:	gap of unknown length
21254	21353:	contig of 1060 bp in length
21354	22737:	gap of unknown length
22738	22837:	contig of 1384 bp in length
22838	23986:	gap of unknown length
23987	23986:	contig of 1049 bp in length
23987	23987:	gap of unknown length
23987	23995:	contig of 1409 bp in length
25346	25495:	gap of unknown length
25496	25906:	contig of 1411 bp in length
26907	27006:	gap of unknown length
27007	28457:	contig of 1451 bp in length
28458	28557:	gap of unknown length
28558	30511:	contig of 1954 bp in length
30512	30611:	gap of unknown length
30612	32038:	contig of 1427 bp in length
32039	32138:	gap of unknown length
32139	33355:	contig of 1217 bp in length
33356	33455:	gap of unknown length
33456	34544:	contig of 1089 bp in length
34545	34644:	gap of unknown length
34645	36300:	contig of 1656 bp in length
36301	36400:	gap of unknown length
36401	38424:	contig of 2024 bp in length
38425	38524:	gap of unknown length
38525	40057:	contig of 1533 bp in length
40058	40157:	gap of unknown length
40158	41398:	contig of 1241 bp in length
41399	41498:	gap of unknown length
41499	42598:	contig of 1040 bp in length
42539	42638:	gap of unknown length

* 42639 44986: contig of 2348 bp in length
* 44987 45086: gap of unknown length
* 45087 46574: contig of 1488 bp in length
* 46575 46674: gap of unknown length
* 46675 48143: contig of 1469 bp in length
* 48144 48243: gap of unknown length
* 48244 49755: contig of 1512 bp in length
* 49756 49855: gap of unknown length
* 49856 50905: contig of 1050 bp in length
* 50906 51006: gap of unknown length
* 51006 52535: contig of 1530 bp in length
* 52536 52635: gap of unknown length
* 52635 54233: contig of 1598 bp in length
* 54234 54333: gap of unknown length
* 54334 55448: contig of 1115 bp in length
* 55449 55548: gap of unknown length
* 55549 56713: contig of 1165 bp in length
* 56714 56813: gap of unknown length
* 56814 58783: contig of 1970 bp in length
* 58784 59667: contig of 1084 bp in length
* 59668 60067: gap of unknown length
* 60068 61165: contig of 1098 bp in length
* 61166 61265: gap of unknown length
* 61266 62383: contig of 1118 bp in length
* 62384 62483: gap of unknown length
* 62484 64426: contig of 1943 bp in length
* 64427 64526: gap of unknown length
* 64527 67010: contig of 2484 bp in length
* 67011 68249: contig of 1139 bp in length
* 68250 68349: gap of unknown length
* 68350 69652: contig of 1303 bp in length
* 69653 69752: gap of unknown length
* 69753 71827: contig of 2075 bp in length
* 71828 71927: gap of unknown length
* 71928 73967: contig of 2040 bp in length

Alignment Scores:

Pred. No.: 1.72 Length: 163087
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-858-332g-15 (1-12) x AC139383 (1-163087)

QY 1 HisanHisAnHisAnHisAnHisAnHisAn 12
Db 150701 CACACCAACCAACCAACCAACCAACCAAC 150666

RESULT 53
AP003356/c 170533 bp DNA linear PRI 28-JUN-2001
LOCUS Homo sapiens genomic DNA, chromosome 8q23, clone: KBI732A1.
DEFINITION AP003356
ACCESSION
VERSION
KEYWORDS
AP003356.2 GI:14572662

SOURCE
ORGANISM
Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

1 (sites)
Shimizu,N. and Asakawa,S.
Homo sapiens DNA chromosome 8 SEQUENCE
Published only in Database (2001)
2 (bases 1 to 170533)
Shimizu,N. and Asakawa,S.
Direct Submission
Submitted (28-FEB-2001) Nobuyoshi Shimizu, Keio University, School
of Medicine, Molecular Biology, 35 Shinanomachi, Shinjuku-Ku, Tokyo
160-8582, Japan (E-mail:nehimizu@med.keio.ac.jp,

COMMENT Tel:81-3-3351-2370, Fax:81-3-3351-2370)
On Jun 27, 2001 this sequence version replaced gi:13359071.
FEATURES
Location/Qualifiers
1..170533

source
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8q23"
/clone="KBI732A1"
/cell_line="FLEB 14 - 14"
/clone_lib="Keio BAC library"
88..124
/evidence=not experimental
repeat_region
/rpt_family="(CA)n"
complement(185..493)
/evidence=not experimental
repeat_region
/rpt_family="AluSg"
557..843
/evidence=not experimental
repeat_region
/rpt_family="MLT2H"
925..1076
/evidence=not experimental
repeat_region
/rpt_family="MLT2H"
1639..1789
/evidence=not experimental
repeat_region
/rpt_family="L2"
2432..2454
/evidence=not experimental
repeat_region
/rpt_family="(TTTA)n"
complement(2455..2732)
/evidence=not experimental
repeat_region
/rpt_family="AluJo"
3010..3115
/evidence=not experimental
repeat_region
/rpt_family="MIR"
3206..3315
/evidence=not experimental
repeat_region
/rpt_family="L2"
3518..3558
/evidence=not experimental
repeat_region
/rpt_family="(CAA)n"
3863..3945
/evidence=not experimental
repeat_region
/rpt_family="L2"
3985..4224
/evidence=not experimental
repeat_region
/rpt_family="MER20"
4337..4371
/evidence=not experimental
repeat_region
/rpt_family="(CA)n"
complement(4418..4718)
/evidence=not experimental
repeat_region
/rpt_family="AluSg"
complement(4719..5018)
/evidence=not experimental
repeat_region
/rpt_family="AluDb"
complement(5089..5139)
/evidence=not experimental
repeat_region
/rpt_family="L2"
5377..5505
/evidence=not experimental
repeat_region
/rpt_family="L2"
complement(6473..6777)
/evidence=not experimental
repeat_region
/rpt_family="AluSx"
complement(7730..8029)
/evidence=not experimental
repeat_region
/rpt_family="AluSg"
8637..8762
/evidence=not experimental
repeat_region
/rpt_family="FLM_A"
complement(9317..9514)
repeat_region

Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kahn,L., Karlas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levin,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McGurt,A., McKernan,K., McPheeters,R.,
Meldrum,D., Menus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Olivier,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rotman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teifaye,S., Theodore,J., Tirrell,A., Travers,W., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 16733)

REFERENCE
AUTHORS
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Biren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Boguslavsky,L., Bouckgeater,B., Brown,A., Burkett,G.,
Camporiano,A., Castle,A., Chospel,Y., Colangelo,W., Collins,S.,
Collins,A., Cooke,P., Dearrellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kahn,L., Karlas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levin,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McGurt,A., McKernan,K., McPheeters,R.,
Meldrum,D., Menus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Olivier,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rotman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teifaye,S., Theodore,J., Tirrell,A., Travers,W., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE
JOURNAL
COMMENT
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submission@genome.wi.mit.edu
----- Project Information
Center project name: L8119
Center clone name: 303_C_4
----- Summary Statistics
Sequencing vector: M13: M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 165466 bases at least Q40
Consensus quality: 171544 bases at least Q30
Consensus quality: 173762 bases at least Q20
Insert size: 18200; agarose-fp
Insert size: 174933; sum-of-contigs
Quality coverage: 4.3 in Q20 bases; agarose-fp
Quality coverage: 4.5 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

* as soon as it is available and the accession number will
* be preserved.
1 1180: contig of 1180 bp in length
* 1181 1280: gap of 100 bp
* 1281 2705: contig of 1425 bp in length
* 2706 2805: gap of 100 bp
* 2806 4387: contig of 1582 bp in length
* 4388 4487: gap of 100 bp
* 4488 6165: contig of 1678 bp in length
* 6166 6265: gap of 100 bp
* 6266 10875: contig of 4610 bp in length
* 10876 10975: gap of 100 bp
* 10976 17071: contig of 6096 bp in length
* 17072 17171: gap of 100 bp
* 17172 23490: contig of 6319 bp in length
* 23491 23590: gap of 100 bp
* 23591 29158: contig of 5568 bp in length
* 29159 29258: gap of 100 bp
* 29259 36785: contig of 7527 bp in length
* 36786 36885: gap of 100 bp
* 36886 44832: contig of 7947 bp in length
* 44833 44932: gap of 100 bp
* 44933 52974: contig of 8042 bp in length
* 52975 53074: gap of 100 bp
* 53075 62479: contig of 9405 bp in length
* 62480 62579: gap of 100 bp
* 62580 71273: contig of 8694 bp in length
* 71274 71373: gap of 100 bp
* 71374 79376: contig of 8003 bp in length
* 79377 79476: gap of 100 bp
* 79477 90216: contig of 10640 bp in length
* 90217 90216: gap of 100 bp
* 90217 105794: contig of 15578 bp in length
* 105795 105894: gap of 100 bp
* 105895 124221: contig of 18327 bp in length
* 124222 124321: gap of 100 bp
* 124322 146798: contig of 22477 bp in length
* 146799 146898: gap of 100 bp
* 146899 176733: contig of 29835 bp in length.
FEATURES
Location/Qualifiers
source
1. 176733
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16"
/clone="RP11-303C4"
/clone_1b="RP11 Human Male BAC"
1. 1180
/note="assembly_fragment"
misc_feature
/note="assembly_fragment"
1281. 2705
/note="assembly_fragment"
misc_feature
/note="assembly_fragment"
2806. 4387
/note="assembly_fragment"
misc_feature
/note="assembly_fragment"
4488. 6165
/note="assembly_fragment"
misc_feature
/note="assembly_fragment"
6266. 10875
/note="assembly_fragment"
misc_feature
/note="assembly_fragment"
10976. 17071
/note="assembly_fragment"
misc_feature
/note="assembly_fragment"
17172. 23490
/note="assembly_fragment"
misc_feature
/note="assembly_fragment"
23591. 29158
/note="assembly_fragment"
misc_feature
/note="assembly_fragment"
29259. 36785
/note="assembly_fragment"
misc_feature
/note="assembly_fragment"
36886. 44832
/note="assembly_fragment"
misc_feature
/note="assembly_fragment"
44933. 52974
/note="assembly_fragment"
misc_feature
/note="assembly_fragment"
53075. 62479
/note="assembly_fragment"
misc_feature
62580. 71273


```

repeat_region      /rpt_family="B2"
24428. .24727
/rpt_family="MAlR"
repeat_region      24729. .24802
/rpt_family="L1"
repeat_region      25431. .25507
/rpt_family="ID"
repeat_region      25514. .25711
/rpt_family="MER1_type"
repeat_region      27264. .27464
/rpt_family="MER1_type"
repeat_region      27747. .28051
/rpt_family="MAlR"
repeat_region      28182. .28467
/rpt_family="B4"
repeat_region      28593. .28971
/rpt_family="L1"
repeat_region      30144. .30479
/rpt_family="L1"
repeat_region      31424. .31643
/rpt_family="L1"
repeat_region      31885. .32259
/rpt_family="MAlR"
repeat_region      32590. .32663
/rpt_family="BCL_MM"
repeat_region      32667. .32779
/rpt_family="Alu"
repeat_region      33647. .33822
/rpt_family="B2"
repeat_region      35418. .35579
/rpt_family="B4"
repeat_region      35692. .36238
/rpt_family="L1"
repeat_region      35887. .35941
/rpt_family="B4"
repeat_region      3742. .39966
/rpt_family="MER1_type"
repeat_region      39967. .40067
/rpt_family="B4"
repeat_region      40035. .40094
/rpt_family="B4"
repeat_region      40132. .40256
/rpt_family="B4"
repeat_region      40737. .41031
/rpt_family="MER1A"
repeat_region      41053. .41227
/rpt_family="L1"
repeat_region      42103. .42298
/rpt_family="B4"
repeat_region      42942. .42972
/rpt_family="MER1_type"
unare             /note="Sequence derived from one plasmid subclone."
repeat_region      43203. .43518
/rpt_family="B4"
repeat_region      44343. .44554
/rpt_family="B4"
repeat_region      45269. .45477
/rpt_family="B2"
repeat_region      45786. .45922
/rpt_family="B2"
repeat_region      45927. .45995
/rpt_family="Alu"
repeat_region      46646. .46792
/rpt_family="B4"
repeat_region      46973. .47231
/rpt_family="Alu"
repeat_region      53665. .53772
/rpt_family="B4"
repeat_region      54230. .54294
/rpt_family="B2"
repeat_region      54300. .54388
/rpt_family="ERV1"
repeat_region      54302. .54413
/rpt_family="Alu"
repeat_region      /rpt_family="B4"

```

```

repeat_region      54566. .54774
/rpt_family="B4"
repeat_region      54936. .54991
/rpt_family="MAlR"
repeat_region      61258. .61600
/rpt_family="MAlR"
repeat_region      61649. .61795
/rpt_family="MIR"
repeat_region      61954. .62184
/rpt_family="MER1_type"
repeat_region      63519. .63592
/rpt_family="B2"
repeat_region      63593. .63718
/rpt_family="B4"
repeat_region      63729. .63837
/rpt_family="Alu"

Alignment Scores:
Pred. No.:      1.9      Length:      179415
Score:          84.00     Matches:      12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              10      Gaps:      0

US-09-858-332G-15 (1-12) x AC139578 (1-179415)

OY      1      H18ASNH18ASNH18ASNH18ASNH18ASN 12
Db      82863 CACACACCAACCAACCAACCAACCAACCAACCAAC 82828

RESULT 56
AC116960
LOCUS      182870 bp      DNA      linear      INV 12-MAR-2003
DEFINITION      Dictyostelium discoideum chromosome 2 map
ACCESSION      AC116960
VERSION      AC116960.2 GI:28850330
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
Gloeckner,G., Eichinger,L., Szafranski,K., Pachbat,J., Dear,P.,
Lehmann,R., Baumgart,C., Parra,G., April,J.F., Guigo,R., Kumpf,K.,
Tunggal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and
Noegel,A.A.
Sequence and analysis of chromosome 2 of Dictyostelium discoideum
Nature 418 (6893), 79-85 (2002)
JOURNAL
MEDLINE
22092622
PUBMED
12097910
REMARK
The Dictyostelium Genome Sequencing Consortium
REFERENCE
2 (bases 1 to 182870)
Baumgart,C.
JOURNAL
Direct Submission
Submitted (04-APR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
3 (bases 1 to 182870)
Baumgart,C.
JOURNAL
Direct Submission
Submitted (05-MAR-2003) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
4 (bases 1 to 182870)
Baumgart,C.
JOURNAL
Direct Submission
Submitted (12-MAR-2003) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
On Mar 5, 2003 this sequence version replaced GI:19920015.
CDS predictions from Genaid do not necessarily reflect true genes.
Further information is available from IMB Jena, Department of
Genome Analysis
(http://genome.imb-jena.de/dictyostelium/)
and the University Cologne, Institute for Biochemistry I

```


Collamore, A., Cook, A., Cooke, P., Corum, B., DeArrellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hages, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menes, L., Mihova, T., Menga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, D., Peterson, K., Phunkhang, P., Pierre, N., Raghupka, A., Ramasamy, U., Raymond, C., Reta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testafaye, S., Theodore, J., Tophan, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, U., Zemek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (17-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 17, 2003 this sequence version replaced gi:2867091.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L23877
Center clone name: 386 N 2
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 182778 bases at least Q40
Consensus quality: 183267 bases at least Q30
Consensus quality: 183499 bases at least Q20
Insert size: 180000; agarose-fp
Insert size: 183702; sum-of-ctnigs
Quality coverage: 9.7 in Q20 bases; agarose-fp
Quality coverage: 9.5 in Q20 bases; sum-of-ctnigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 14629: contig of 14629 bp in length
* 14630 14729: gap of 100 bp
* 14730 16226: contig of 1497 bp in length
* 16227 16326: gap of 100 bp
* 16327 70387: contig of 54061 bp in length
* 70388 70487: gap of 100 bp
* 70488 125864: contig of 53377 bp in length
* 125865 125964: gap of 100 bp
* 125965 180752: contig of 54788 bp in length
* 180753 180852: gap of 100 bp
* 180853 184202: contig of 3350 bp in length.
Location/Qualifiers
1. 184202
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-386N2"
/clone_lib="RRC1-23 Female Mouse BAC"
1. 14629
/note="assembly_fragment"
clone_end:SP6

vector side:left"
14730..16226
/note="assembly_fragment"
misc_feature
16327..70387
/note="assembly_fragment"
misc_feature
70488..125864
/note="assembly_fragment"
misc_feature
125965..180752
/note="assembly_fragment"
180853..184202
/note="assembly_fragment"
clone_end:T7
vector_side:right"

ORIGIN

Alignment Scores:
Pred. No.: 1.95 Length: 184202
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-858-332G-15 (1-12) x AC116475 (1-184202)

Oy 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 121052 CATRACCATTAACCATTAACCATTAACCATTAACCATTAAC 121087

RESULT 59

AC011405/c

LOCUS

DEFINITION

AC011405

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NOTE: This is a 'working draft' sequence. It currently

consists of 6 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 14629: contig of 14629 bp in length

* 14630 14729: gap of 100 bp

* 14730 16226: contig of 1497 bp in length

* 16227 16326: gap of 100 bp

* 16327 70387: contig of 54061 bp in length

* 70388 70487: gap of 100 bp

* 70488 125864: contig of 53377 bp in length

* 125865 125964: gap of 100 bp

* 125965 180752: contig of 54788 bp in length

* 180753 180852: gap of 100 bp

* 180853 184202: contig of 3350 bp in length.

Location/Qualifiers

1. 184202

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone="RP23-386N2"

/clone_lib="RRC1-23 Female Mouse BAC"

1. 14629

/note="assembly_fragment"

clone_end:SP6

FEATURES

source

1. 184541

/organism="Homo sapiens"

location/Qualifiers

1. 184541

/organism="Homo sapiens"

```
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTB-46B19"

ORIGIN
Alignment Scores:
Pred. No.: 1.95 Length: 184541
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-858-332G-15 (1-12) x AC011405 (1-184541)

Oy 1 HisAenh1sAenh1sAenh1sAenh1sAenh1sAen 12
Db 27855 CATTAATCATTAATCATTAATCATTAATCATTAAT 27820

RESULT 60
AC131664 188388 bp DNA linear HTG 13-MAY-2003
LOCUS AC131664/c
DEFINITION Mus musculus chromosome UNK clone RP23-274J16, WORKING DRAFT
ACCESSION AC131664
VERSION AC131664.2 GI:30581677
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 188388)
McPherson,J.D. and Waterston,R.H.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 188388)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (13-MAY-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On May 13, 2003 this sequence version replaced gi:22475537.

REFERENCE
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
COMMENT Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@genome.wustl.edu
Project name: M_BA0274J16
Center project name: M_BA0274J16

----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: Plasmid; 100%
Chemistry: Dye-terminator; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 186216 bases at least Q40
Consensus quality: 18652 bases at least Q30
Consensus quality: 186941 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 5628: contig of 5628 bp in length
* 5628 5728: gap of unknown length
* 5729 23196: contig of 17468 bp in length
* 23197 23296: gap of unknown length
* 23297 50200: contig of 26904 bp in length
* 50201 50300: gap of unknown length
* 50301 107264: contig of 56864 bp in length
* 107265 107365: gap of unknown length
* 107365 188388: contig of 81024 bp in length.
Location/Qualifiers
1..188388
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-274J16"
1..5628
/feature="assembly_name:Contig16"
5729..23196
/feature="assembly_name:Contig17"
23297..50200
/feature="assembly_name:Contig18"
50301..107264
/feature="assembly_name:Contig19"
107365..188388
/feature="assembly_name:Contig20"

ORIGIN
Alignment Scores:
Pred. No.: 2 Length: 188388
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-858-332G-15 (1-12) x AC131664 (1-188388)

Oy 1 HisAenh1sAenh1sAenh1sAenh1sAenh1sAen 12
Db 8183 CATTAACCATTAACCATTAACCATTAACCATTAAC 8148

RESULT 61
AC123741 188453 bp DNA linear ROD 09-JUL-2004
LOCUS AC123741/c
DEFINITION Mus musculus chromosome 1, clone RP24-113A5, complete sequence.
ACCESSION AC123741
VERSION AC123741.9 GI:50080382
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 188453)
Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukigalter,B., Brown,A., Camarata,J., Campiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Dewar,K., Diaz,J.S., Dodge,S.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N.,
Mathews,C., McCarthy,M., McEwan,P., McKernan,K., Melidrim,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
```

TITLE
JOURNAL
REFERENCE
AUTHORS

O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Ribbeck, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, K., Sudriaman, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 188453)

TITLE
JOURNAL
REFERENCE
AUTHORS

Birren, B., Nusbaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H. M., Banna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Cornu, B., DeArliano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Mathews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (07-JUL-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 188453)

TITLE
JOURNAL
REFERENCE
AUTHORS

Birren, B., Nusbaum, C., Lander, E., Aboueleil, A., Allen, N., Bloom, T., Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Cornu, B., DeArliano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Mathews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (09-JUL-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 9, 2004 this sequence version replaced gi:49745121.

COMMENT
 All repeats were identified using RepeatMasker:

http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@broad.mit.edu
 ----- Project Information
 Center project name: L26529
 Center clone name: 113_A_5

FEATURES

 source
 1..188453
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="1"
 /map="1"
 /clone_11b="RP24-113A5"
 /clone_11b="RP24-113A5"
 1..4
 /note="clone boundary
 clone end:SP6
 site:MboI"
 complement(210..697)
 /rpt_family="L1_MM"
 complement(3240..4210)
 /rpt_family="Lx8"
 4430..4474
 /rpt_family="CA)n"
 4756..4796
 /rpt_family="TG)n"
 5336..5378
 /rpt_family="AT_rich"
 5827..6424
 /rpt_family="L1_MM"
 complement(5962..6472)
 /note="single clone coverage"
 6425..6451
 /rpt_family="AT_rich"
 complement(7566..8612)
 /rpt_family="Lx7"
 complement(8615..9281)
 /rpt_family="Lx5"
 9282..9322
 /rpt_family="TG)n"
 complement(9323..10080)
 /rpt_family="Lx5"
 complement(10081..10223)
 /rpt_family="RMER1B"
 complement(10590..10672)
 /rpt_family="RMER1B"
 complement(10977..11251)
 /rpt_family="RMER1B"
 complement(11306..11765)
 /rpt_family="Lx5"
 complement(11766..12110)
 /rpt_family="ORR1C"
 complement(12118..12269)
 /rpt_family="Lx5"
 complement(12250..12437)
 /rpt_family="Lx5"
 complement(12418..12554)
 /rpt_family="L1"
 12555..12590
 /rpt_family="TG)n"
 12590..12650
 /rpt_family="GA)n"
 complement(12591..12704)
 /rpt_family="L1"
 12984..13017
 /rpt_family="AT_rich"
 13225..13287
 /rpt_family="TRATG)n"
 13582..13620
 /rpt_family="CA)n"
 13632..14105
 /rpt_family="RLR1A1A"
 14110..14158
 /rpt_family="TG)n"
 complement(14277..14340)
 /rpt_family="ID4"
 complement(14515..14888)


```

repeat_region /rpt_family="MTD"
15541..15573
/rpt_family="TG)n"
repeat_region complement(15595..15681)
/rpt_family="RSIN1"
repeat_region complement(15849..16220)
/rpt_family="MLT1A"
repeat_region complement(16331..16402)
/rpt_family="U5"
repeat_region 16403..16513
/rpt_family="LX5"
repeat_region 18093..18113
/rpt_family="AT_rich"
repeat_region complement(18359..18462)
/rpt_family="LX9"
repeat_region 18463..18504
/rpt_family="AT_rich"
repeat_region complement(18935..19149)
/rpt_family="L1F"
repeat_region 19153..20136
/rpt_family="L1_MM"
repeat_region 21300..21644
/rpt_family="OR1D"
repeat_region 21823..21860
/rpt_family="AT_rich"
repeat_region 22120..22406

```

Alignment Scores:

Pred. No.:	2	Length:	188453
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-858-332g-15 (1-12) x AC123741 (1-188453)

Qy 1 HisAnHISAenHISAenHISAenHISAen 12
 Db 12558 CACACCAACCAACCAACCAACCAACCAACCAAC 12523

RESULT 62
 CR450684/c 189105 bp DNA linear HTG 27-MAY-2004
 LOCUS Danio rerio clone DKEY-60D5, WORKING DRAFT SEQUENCE, 3 unordered
 DEFINITION pieces.
 ACCESSION CR450684
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS Danio rerio (zebrafish)
 SOURCE Danio rerio
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 189105)
 Direct Submission
 Submitted (25-MAY-2004) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On May 27, 2004 this sequence version replaced gi:47604342.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: zfish-help@sanger.ac.uk
 ----- Project Information
 Center project name: ZK60D5
 ----- Summary Statistics
 Assembly Program: XGAP4; version 4.5
 Chemistry: Dye-terminator; 100% of reads
 Consensus quality: 187965 bases at least Q40
 Consensus quality: 188193 bases at least Q30

Consensus quality: 188315 bases at least Q20
 Insert size: 188905; sum-of-contigs
 Insert size: 178551; 2.8% error; agarose-fp
 Quality coverage: 10..11x in Q20 bases; sum-of-contigs Quality
 coverage: 10..97x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available with the accession number will
 * be preserved.

FEATURES
 source 1..22085: contig of 22085 bp in length
 * 22086 22185: gap of 100 bp
 * 22186 101934: contig of 79749 bp in length
 * 101935 102034: gap of 100 bp
 * 102035 189105: contig of 87071 bp in length.
 Location/Qualifiers
 1..189105
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="DKEY-60D5"
 /clone_11b="DanioKey"

misc_feature 1..22085
 /note="assembly_fragment:00038.0"
 misc_feature 22186..101934
 /note="assembly_fragment:00421"
 misc_feature 102035..189105
 /note="assembly_fragment:01537"

ORIGIN

Alignment Scores:

Pred. No.:	2	Length:	189105
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-858-332g-15 (1-12) x CR450684 (1-189105)

Qy 1 HisAnHISAenHISAenHISAenHISAen 12
 Db 158034 CATATCATATCATATCATATCATATCATAT 157999

RESULT 63
 AC132110/c 189130 bp DNA linear ROD 06-FEB-2004
 LOCUS Mus musculus BAC clone RP24-188M9 from chromosome 18, complete
 DEFINITION sequence.
 ACCESSION AC132110
 VERSION AC132110.2 GI:41386975
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 189130)
 Van Brunt A., Haakenson, W., Van Brunt, A., Kozlowski, A. and
 Bielicki, L.
 The sequence of Mus musculus BAC clone RP24-188M9
 Unpublished (2001)
 2 (bases 1 to 189130)
 Wilson, R.
 Sequencing of Mus musculus
 Unpublished (2001)
 3 (bases 1 to 189130)
 McPherson, J.D. and Waterston, R.H.
 Direct Submission
 Submitted (29-AUG-2002) Genome Sequencing Center, 4444 Forest Park

REFERENCE
 AUTHORS Van Brunt A., Haakenson, W., Van Brunt, A., Kozlowski, A. and
 Bielicki, L.
 TITLE The sequence of Mus musculus BAC clone RP24-188M9
 JOURNAL Unpublished (2001)
 REFERENCE
 AUTHORS Wilson, R.
 TITLE Sequencing of Mus musculus
 JOURNAL Unpublished (2001)
 REFERENCE
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (29-AUG-2002) Genome Sequencing Center, 4444 Forest Park

REFERENCE
AUTHORS
TITLE
JOURNAL
JOURNAL
AUTHORS
TITLE
JOURNAL
COMMENT

Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 189130)
Wilson, R.K.
Direct Submission
Submitted (29-JUN-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 189130)
Wilson, R.
Direct Submission
Submitted (06-FEB-2004) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jan 29, 2004 this sequence version replaced gi:22539272.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@wustl.wustl.edu
----- Summary Statistics
Center project name: M_B018BM09

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence, and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:
The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.choi.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.choi.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone. This clone is overlapped by AC122826

FEATURES

source

Location/Qualifiers
1..189130
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="18"
/map="18"
/clone="RP24-188M9"
/clone_id="RPCI-24"
45..281
repeat_region
/rpt_family="MaLR"
337..554
/rpt_family="MaLR"
693..926
/rpt_family="L1"
930..1298
/rpt_family="MaLR"
3943..4217
/rpt_family="L1"
4973..5118
/rpt_family="Alu"
5934..6208
repeat_region

/rpt_family="B4"
6333..6575
/rpt_family="MaLR"
8955..9233
repeat_region
/rpt_family="B4"
10160..10306
/rpt_family="Alu"
10603..10793
repeat_region
/rpt_family="B4"
11606..11671
/rpt_family="ID"
12296..12526
repeat_region
/rpt_family="B4"
12556..12789
repeat_region
/rpt_family="MaLR"
13320..14098
repeat_region
/rpt_family="B4"
14188..14347
repeat_region
/rpt_family="B4"
15090..15137
repeat_region
/rpt_family="B4"
16274..16418
repeat_region
/rpt_family="MER1_type"
16526..16784
repeat_region
/rpt_family="B4"
16761..16844
repeat_region
/rpt_family="ERVK"
16845..16895
repeat_region
/rpt_family="ERV1"
16867..16918
repeat_region
/rpt_family="ERV1"
17615..17876
repeat_region
/rpt_family="B4"
18353..18428
repeat_region
/rpt_family="ID"
20048..20159
repeat_region
/rpt_family="Alu"
20524..20607
repeat_region
/rpt_family="MER2_type"
20670..20804
repeat_region
/rpt_family="Alu"
20809..20883
repeat_region
/rpt_family="ID"
21026..21206
repeat_region
/rpt_family="B2"
21650..21768
repeat_region
/rpt_family="L1"
23458..23843
repeat_region
/rpt_family="MaLR"
24068..24200
repeat_region
/rpt_family="MaLR"
24201..24624
repeat_region
/rpt_family="ERVK"
24625..24846
repeat_region
/rpt_family="MaLR"
24905..25284
repeat_region
/rpt_family="MaLR"
26075..26216
repeat_region
/rpt_family="B4"
26237..26420
repeat_region
/rpt_family="B4"
26438..26514
repeat_region
/rpt_family="ID"
26612..26744
repeat_region
/rpt_family="MaLR"
27909..27978
repeat_region
/rpt_family="ID"
28166..28297
repeat_region
/rpt_family="B4"
28307..28546
repeat_region
/rpt_family="L1"
28543..28670
repeat_region
/rpt_family="L1"

```

repeat_region      28677..28727
                    /rpt_family="B4"
repeat_region      29224..29365
                    /rpt_family="Alu"
repeat_region      30617..30734
                    /rpt_family="Alu"
repeat_region      32171..32227
                    /rpt_family="ERV1"
repeat_region      32445..32601
                    /rpt_family="B4"
repeat_region      32667..32798
                    /rpt_family="B4"
repeat_region      32880..32997
                    /rpt_family="B2"
repeat_region      33034..33177
                    /rpt_family="B4"
repeat_region      33118..33227
                    /rpt_family="Alu"
repeat_region      33597..33674
                    /rpt_family="B4"
repeat_region      34429..34495
                    /rpt_family="ERV1"
repeat_region      36865..36944
                    /rpt_family="ID"
repeat_region      37114..37262
                    /rpt_family="B4"
repeat_region      37725..37895
                    /rpt_family="B4"
repeat_region      38211..38327
                    /rpt_family="Alu"
repeat_region      38356..38420

```

Alignment Scores:

```

Pred. No.:      2      Length:      189130
Score:          84.00   Matches:      12
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match:      100.00%   Indels: 0
DB:              10      Gaps: 0

```

US-09-858-332g-15 (1-12) x AC132110 (1-189130)

Qy 1 H18A8nH1sA8nH1sA8nH1sA8nH1sA8n 12

Db 16658 CACACCAATACACACACACACACACACACACAC 16623

RESULT 64

AC134540

LOCUS Mus musculus BAC clone RP24-400M10 from chromosome 18, complete

DEFINITION sequence.

AC134540

VERSION AC134540.3 GI:45237306

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 189505)

AUTHORS Ali J., Bielicki L., Meyer R. and Haekenson W.

TITLE The sequence of Mus musculus BAC clone RP24-400M10

JOURNAL Unpublished (2001)

REFERENCE 2 (bases 1 to 189505)

AUTHORS McPherson J.D. and Waterston R.H.

TITLE Direct Submission

JOURNAL Submitted (27-SEP-2002) Genome Sequencing Center, 4444 Forest Park

REFERENCE 3 (bases 1 to 189505)

AUTHORS McPherson J.D. and Waterston R.H.

TITLE Direct Submission

JOURNAL Submitted (04-OCT-2002) Genome Sequencing Center, 4444 Forest Park

REFERENCE 4 (bases 1 to 189505)

AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-2004) Genome Sequencing Center, 4444 Forest Park
REFERENCE 5 (bases 1 to 189505)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park
COMMENT On Mar 6, 2004 this sequence version replaced gi:23499622.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUSC
Web site: <http://genome.wustl.edu>
Contact: submissions@watson.wustl.edu
----- Summary Statistics
Center project name: M_BB0400M10

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RP24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC132110.

FEATURES

source

Location/Qualifiers

1..189505

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="18"

/map="18"

/clone="RP24-400M10"

/clone_1lb="RP24-24"

/7124..77476

/note="Sequence derived from PCR product of project DNA."

78150..78171

/note="Sequence derived from one plasmid subclone."

90757..90881

/note="Unresolved simple sequence repeat."

118118..118132

/note="Sequence derived from one plasmid subclone."

ORIGIN

Alignment Scores:

```

Pred. No.:      2.01      Length:      189505
Score:          84.00   Matches:      12
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match:      100.00%   Indels: 0
DB:              10      Gaps: 0

```


FNNEVLIARLQHLNVLRLACCVDAGEKMLIYEYLLENLSLDHLEFDSRSKLNQM
RFDIINGIARGLLYLHODSRFRIIHRDLKASWILLDKYMTPTKISFGMARLFGDETE
AATRKVYGVGYMSPEYAMDGI FSKMSDVSFSGVTLLEIISKENKGYNDRLNL
GCVWNNMKSGKLEIIDIPTTSSSTFROHEILRQIQLLCVDERADRPMTSLVILL
MGSESTTIPQKAPGYCLERSLDDTSSSSKORDDESMVTWQITVSVLJAR"
complement (7083. .7409)
/gene="AT4g21380"
/number=1
complement (7410. .7493)
/gene="AT4g21380"
/number=1
complement (7494. .7644)
/gene="AT4g21380"
/number=2
complement (7645. .7740)
/gene="AT4g21380"
/number=2
complement (7741. .7978)
/gene="AT4g21380"
/number=3
complement (7979. .8069)
/gene="AT4g21380"
/number=3
complement (8070. .8280)
/gene="AT4g21380"
/number=4
complement (8281. .8358)
/gene="AT4g21380"
/number=4
complement (8359. .8552)
/gene="AT4g21380"
/number=5
complement (8553. .8641)
/gene="AT4g21380"
/number=5
complement (8642. .8776)
/gene="AT4g21380"
/number=6
complement (8777. .9657)
/gene="AT4g21380"
/number=6
complement (9658. .10954)
/gene="AT4g21380"
/number=7
/gene="AT4g21390"
complement (join(12322. .12639,12758. .12905,13005. .13242,
13324. .13534,13619. .13854,13940. .15338))
/gene="AT4g21390"
complement (join(12322. .12639,12758. .12905,13005. .13242,
13324. .13534,13619. .13854,13940. .15338))
/gene="AT4g21390"
/note="strong similarity to serine/threonine kinase BRK,
Brassica oleracea, gb:Y12531
Contains prokaryotic membrane lipoprotein lipid attachment
site AA392-402;Protein kinases signatures and profile
AA650-662
contains EST gb:T04455, T43643, AI992665.1, T41816,
AI996080.1, T21059, AI996706.1, AU036593.1, T21145,
R65507, AI993345.1"
/codon_start=1
/product="serine/threonine kinase-like protein"
/protein_id="CA81246.1"
/db_xref="GI:7268936"
/db_xref="GOA:O81906"
/translation="MRFRKTSIVLSLPLVFESSMAANTIRGSLDNGIHKPL
VSPKTFELGPGSSGTHRLPGITWCIIEKAYVWVANRATPISDSGVLMISNDN
LVLDGKNITWSSNIESSTNNNNRVSIHDTGNFVLSSTDRPIWSEFNPTDFF
LPQKRVYRNPOTGDHAFVSWRSETDPSPGVSLGVDPGAPETIWMGKTKRMRG
QMSAIPTGIPNMSLLTNVLYGFKLSPPDETSVPTVYSPSVLLRPVLYNGTE
BELRWNETLKMTKFOSEPDSECDQYNRGKFGICDMKSGNGICSCIHGYEYVSGNW
SRGGRRTPLCERNISVGEDEFLTKSVKLPDEIPEHNIVDEPDERERCINRCSGN

AVSLVGGICMWNODLVLDQOFBAGSSLIHRLADSENGENRKTIAIVAVYGV
LIIGIFALLMRFRKXDDVGAGCGNNTDTSVVADLTYSKETTSFSSGVIMIGKA
VNTSELFPVSNIAIAIATNDPKENELRGCGRPYKGLNEDGRITAVKRSKSGQG
VBEFKNEIILIAKIQHRNIVRLGCGFBEERKMLVYEWPNKSLDFLEDETKALID
WKLRSTIIGIARGLLYLHODSRRLRIIHRDLKASVVLDAENPKISDPFGARIPGN
QNSANTVRVGVGYMSPEYAMEGLFSVKSIVSFGVLLLEIVSGKRTSLRSSPHS
LIGYAVLYTVHGRSESLVDPKIRVTSKREALRCIHVAMLCVDSAAERPMASVLLM
LSSDPTATLAPRQPTFTSTRNSIDVNFALDSQQQYIVSSNEIISTVVLGR"
complement (12322. .12639)
/gene="AT4g21390"
/number=1
complement (12640. .12757)
/gene="AT4g21390"
/number=1
complement (12758. .12905)
/gene="AT4g21390"
/number=2
complement (12906. .13004)
/gene="AT4g21390"
/number=2
complement (13005. .13242)
/gene="AT4g21390"
/number=3
complement (13243. .13323)
/gene="AT4g21390"
/number=3
complement (13324. .13534)
/gene="AT4g21390"
/number=4
complement (13535. .13618)
/gene="AT4g21390"
/number=4
complement (13619. .13854)
/gene="AT4g21390"
/number=5
complement (13855. .13939)
/gene="AT4g21390"
/number=5
complement (13940. .15338)
/gene="AT4g21390"
/number=6
/gene="AT4g21400"
complement (join(17082. .17390,17476. .17632,17702. .18240,
18325. .18437,18489. .18665,18733. .19573))
/gene="AT4g21400"
complement (join(17082. .17390,17476. .17632,17702. .18240,
18325. .18437,18489. .18665,18733. .19573))
/gene="AT4g21400"
/note="similarity to receptor protein kinase (IRK1),
Ipomoea trifida, gb:U20948
Contains protein kinases signatures and profile
AA367-389;Protein kinases signatures and profile AA510-522
contains EST gb:AI996706.1"
/codon_start=1

Alignment Scores:
Pred. No.: 2.07 Length: 194916
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-858-332g-15 (1-12) x ATCHRIV55 (1-194916)

Qy 1 H1SAenH1SAenH1SAenH1SAenH1SAen 12
|||||
Db 194880 CACAATCACAATCACAATCACAATCACAAT 194915
|||||

RESULT 67
ATCHRIV56
LOCUS ATCHRIV56 196247 bp DNA linear PLN 16-MAR-2000

DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 56.
ACCESSION AL161556
VERSION AL161556.2 GI:7269026
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
JOURNAL
REFERENCE
AUTHORS
JOURNAL
AUTHORS
MURPHY,G., RIDLEY,P., HUDSON,S., MEWES,H.W., LEMCKE,K. and
MAYER,K.F.X.
Unpublished
2 (bases 157343 to 167411)
Robben,J., GRYMONPREZ,B., VOLCKAERT,G., MEWES,H.W., LEMCKE,K. and
MAYER,K.F.X.
Unpublished
3 (bases 167197 to 196247)
MURPHY,G., RIDLEY,P., HUDSON,S., MEWES,H.W., LEMCKE,K. and
MAYER,K.F.X.
Unpublished
4 (bases 1 to 196247)
EU Arabidopsis sequencing, project.
Direct Submission
Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project
Coordinator: Mike Bavan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bavan@birc.ac.uk
Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>
this fragment has an overlap with ATCHRIV55 at the 5' end and an
overlap with ATCHRIV57 at the 3' end.
Location/Qualifiers
1. 196247
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
5510..11823
/gene="AT4g21820"
join(5510..6217,6310..6778,7223..7449,7548..7613,
7799..7852,8220..8336,8544..8642,8699..8879,9087..9224,
9426..9620,9702..9982,10872..11094,11161..11425,
11469..11624,11736..11823)
/note="similarity to calmodulin-binding protein SHA1, Mus
musculus, AF062378
contains EST gb:AI997648.1"
/codon_start=1
/product="putative protein"
/protein_id="CAB79137.1"
/db_xref="GI:7269027"
/db_xref="UniProt/TREMBL:O49705"
/translation="MDENPPQSPAPPPNPPASLLSDISNFTKPTSTVSNSTSKS
PYRHPFTAKQTPKSSSNFRPSMWHASRSKVTSSRLKAFELQSGSRKREL
TEKNIKSLAKSLTVAFLFENPENGCDPEFNEGVGNLGRKRDSEALGNSKSL
GVDTWMSPKRLNLGWCGRKSEIDSSLTGSKYLTLESILDVCSLDLQKRMQFL
LSIGCKEIPDVTWTRVSKNIDEGRIKMPQCPVTFGMEKAIKALMEYNQWMLG
LYIIFGSDFLSDSEVNSDOEMAFLLVVISKQPSHDGLARSVYNKVEGYLRPGY
BALGVNLIKRLILLVVIDRKSQSLSTKYGDGIDGSGPLMFSKSKSKSHOLIC
GSEFDMRLVLIKQKVVPSDNRKONANCRILQYKDAQVSLKDDGMMITGEDVAD
GRELITSLIMNI FVHLQPLINGRLILBEIYIKVQGVDEONQITMSTPLEMLNIQ
DRGGEGQSVMSNTDYHDAVONFLLSQKLTALGSPFELGDLLENNAVSNQSVIL
LAFLSKLIIVKENLDPFTLAGYLTNSVFSOEPLNFKLCSQCOVEKYSRIKSCS
SESAVTIEEPDRENGEDATRRFOAIRAMQDMANQOISVGRKNSHTLQSGSRKCTY
DSORAYLLAEIAAVIIQSNIRGLHARRKPRNRRAICFLOAARTMLSVKNIQVEX
FNEEYVTLHLSERSANLKPARVYVKFIVDSRPRIKRSVSVIOZKAVRRHOSNLHET
KALARIQLMRSYKEKVISITIOSYVRGHTRRNRRTYFSSILIQFOVSPFPYISQ
SHMRGYLTRKASQAVLDLRVMQTSANIIDDKRLINKLSLSELSLSELMKVHNILH
ICTETESATKYSDBKCEBELVAAIGAILKLTLLRSASRSIPDOQVSKHALSTIGHARY

POMADELINTKSGISQITFWELLRSFSSQVFLSPKKNKEAYFIASDVIKKICSSHG
VEAVVRKLPLAVRLHVLVEELTRKANIERKNVKGQTKESERRLKEAIELVLTTRR
"

5510..6217
/gene="AT4g21820"
/number=1
6218..6309
/gene="AT4g21820"
/number=1
6310..6778
/gene="AT4g21820"
/number=2
6779..7222
/gene="AT4g21820"
/number=2
7223..7449
/gene="AT4g21820"
/number=3
7450..7547
/gene="AT4g21820"
/number=3
7548..7613
/gene="AT4g21820"
/number=4
7614..7798
/gene="AT4g21820"
/number=4
7799..7852
/gene="AT4g21820"
/number=5
7853..8219
/gene="AT4g21820"
/number=5
8220..8336
/gene="AT4g21820"
/number=6
8337..8543
/gene="AT4g21820"
/number=6
8544..8642
/gene="AT4g21820"
/number=7
8643..8698
/gene="AT4g21820"
/number=7
8699..8879
/gene="AT4g21820"
/number=8
8880..9086
/gene="AT4g21820"
/number=8
9087..9224
/gene="AT4g21820"
/number=9
9225..9425
/gene="AT4g21820"
/number=9
9426..9620
/gene="AT4g21820"
/number=10
9621..9701
/gene="AT4g21820"
/number=10
9702..9982
/gene="AT4g21820"
/number=11
9983..10871
/gene="AT4g21820"
/number=11
10872..11094
/gene="AT4g21820"
/number=12
11095..11160

exon	/gene="AT4g21820" /number=12 11161. .11425 /gene="AT4g21820" /number=13 11426. .11468 /gene="AT4g21820" /number=13 11469. .11624 /gene="AT4g21820" /number=14 11625. .11735 /gene="AT4g21820" /number=14 11736. .11823 /gene="AT4g21820" /number=15 12517. .13199 /gene="AT4g21830" complement(join(12517. .12675,12834. .12990,13090. .13199) /gene="AT4g21830" complement(join(12517. .12675,12834. .12990,13090. .13199) /gene="AT4g21830" /note="strong similarity to CG1-131 protein, Homo sapiens, AF151889 contains EST gb:N38012, AA713335, A1996517.1, AA712490, H67478" /codon_start=1 /product="putative protein" /protein_id="CAB79138.1" /db_xref="GI:7269028" /db_xref="UniProt/TpEMBL:O49706" /translation="MTAAVPAATGSFQKDEWFAVISPFQVRLKLTGTDKRGKEF TKKPEGNYTSCAGCTALYKSTTKFDSGCCMPAFDAIPCAIKQTPEDAGRMETCA VVDGHLGHVFKKEGISTPTDQRHCVNSVSLKFSAGSSQ" complement(12517. .12675) /gene="AT4g21830" /number=1 complement(12675. .12833) /gene="AT4g21830" /number=1 complement(12834. .12990) /gene="AT4g21830" /number=2 complement(12991. .13089) /gene="AT4g21830" /number=2 complement(13090. .13199) /gene="AT4g21830" /number=3 15183. .16130 /gene="AT4g21840" complement(join(15183. .15341,15764. .15920,16015. .16130))
intron	
exon	
intron	
exon	
gene	
CDS	

Alignment Scores:	
Seq. No.: 2.08	Length: 196247
Percent Similarity: 84.00	Matches: 12
Local Similarity: 100.00%	Conservative: 0
Local Similarity: 100.00%	Mismatches: 0
Match: 100.00%	Indels: 0
	Gaps: 0

Page: 0
9-858-333G-15 (1-12) x ATCHRIV56 (1-196247)

```

1  H1SAsnH1sAsnH1sAsnH1sAsnH1sAsn 12
   |||||
4964 CACATCACAATCACAATCACAATCACAAT 4999

```

T	68
:863/c	
AC104d63	198321 bp
Mus musculus chromosome 6,	DNA linear ROD 30-SEP-2003
clone RP23-109C11,	complete sequence.

VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
FORMAT

BOOKNAME
REFERENCE
AUTHORS

TITLE
JOURNAL

REFERENCE

ACI04863.12 GI:30141979
HTG.
Mus musculus (house mouse)
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 198321)

Britten, D., Nusbaum, C. and Lander, E.
Mus musculus chromosome 6, clone RP23-109C11
Unpublished
2 (bases 1 to 198321)
Britten, B., Linton, L., Nusbaum, C., Lander, E., Aiz, A., Allen, N.,
Anderson, S., Berra, N., Brown, J., Buxton, V., Brown, J.

Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazarro, B., Choepel, Y., Colangelo, M., Collins, S., Collimore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Fogarty, D., Gogoravsky, L., Soukngateer, B.

Arreola, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Glnde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, C., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lapocchia, K.

Lamazaes, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrim, J., Menzies, I., Wibaux, T.

Notbu, D., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliveira, L., Oliveira, J., Mieriga, V., Murphy, T., Naylor, J., Nguyen, C.,
Oliver, J., Peterson, K., Plunkharg, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Ryland, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.

Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severly, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Godham, K., Travers, M., Travis, N.

Vieli, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (2-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 198321)
Birren, B., Nusbaum, C., Lander, E. & Altshuler, D.

Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collimore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,

Faro, S., Erickson, J., Dorflis, L., Booley, K., Moore, S., Pitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, P., Jones, C., Kozlov, A., Leung, Y., Liu, X., Macdonald, R., Maruyama, T., McQuinn, K., Miller, E., Morrell, A., Murray, C., Narasimhan, B., O'Connell, J., O'Donnell, J., O'Hara, T., O'Sullivan, T., Padilla, N., Panesar, F., Perrele, P., Pillemer, D., Quinn, J., Reuter, M., Riechers, D., Roediger, J., Rowley, N., Salazar, V., Sarin, A., Schmitt, H., Sheth, S., Smith, J., Sothmann, B., Stach, K., Stein, W., Steiner, J., Strickland, D., Sun, M., Swartz, D., Taniguchi, M., Taylor, A., Thaler, A., Thompson, T., Truemper, J., Turner, M., Urry, M., Valiante, R., Veale, S., Verstra, A., Villarreal, B., Virani, S., Wallack, J., Webb, J., Willott, C., Woodward, C., Wright, F., Wu, K., Xu, S., Ziegler, G.

Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
McDermid, T., Morozov, T., Milnes, E., et al.

Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunghang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rooy, P.

Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Ward, D., Williams, J., and Wright, J.

Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome

Anderson, S., Arachchi, H.M., Barna N., Bastien V., Allen, N.,
Blair, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Cresswell, J., 320 Chaires Street, Cambridge, MA 02141, USA
(bases 1 to 198321)

Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K., Diaz, J. S., Dodge, S., Doolley, K., Dorris, L., Erickson, J., Faro, S.,

Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R.,
Maddali, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Karaman, M., Katsaridis, N., Gage, D., Galagan, J., Gardyna, S.,
Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

Linblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meness, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,

REFERENCE	Nguyen,C., Micol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunhthang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupbach,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
JOURNAL	Submitted (26-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS	5 (bases 1 to 198321)
REFERENCE	Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Arachchi,H.M., Batra,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,S., Chang,J., Choepel,Y., Collamore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gerhart,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hages,B., Hall,J., Horcon,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamae,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,X., Mabbitt,R., McClean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Medlirin,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunhthang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupbach,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
JOURNAL	Submitted (30-SEP-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Apr 26, 2003 this sequence version replaced gi:3065658. All repeats were identified using RepeatMasker: Smt, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html
FEATURES	----- Genome Center Center: Whitehead Institute/ MIT Center for Genome Research Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu ----- Project Information Center project name: L16460 Center clone name: 109_C.11 ----- Location/Qualifiers 1. 198321 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" /chromosome="6" /map="6" /clone="RP23-109C11" /clone_lib="RPCT-23 Female Mouse BAC" 1. 6 /note="Clone boundary clone end:SP6 end site:ECORI" repeat_region complement(21..818) /rpt_family="L1" 820..7113 /rpt_family="L1_MM" 2599..2626 /note="single clone coverage" 2663 /note="Probably T, possibly TT" complement(7115..10490) /rpt_family="L1" 11031..11149 /rpt_family="CT-rich"

	repeat_region	11530..11586	/rpt_family="(TAGA)n"
	repeat_region	11749..11776	/rpt_family="AT_rich"
	repeat_region	12863..12926	/rpt_family="MIR"
	repeat_region	14087..14121	/rpt_family="(TG)n"
	repeat_region	complement(14218..14383)	/rpt_family="B4A"
	repeat_region	15823..16059	/rpt_family="B4A"
	repeat_region	complement(16475..16765)	/rpt_family="RMER19B"
	repeat_region	complement(16883..17078)	/rpt_family="L1_MM"
	repeat_region	complement(17079..17461)	/rpt_family="RMER19B"
	repeat_region	18280..18350	/rpt_family="ORR1C"
	repeat_region	18484..18580	/rpt_family="MTE"
	repeat_region	18586..18640	/rpt_family="(TG)n"
	repeat_region	18927..19362	/rpt_family="RMER6A"
	repeat_region	19623..20015	/rpt_family="L1M4"
	repeat_region	20900..20952	/rpt_family="(TG)n"
	repeat_region	21205..21256	/rpt_family="ID_B1"
	repeat_region	23070..23097	/rpt_family="AT_rich"
	repeat_region	23963..23986	/rpt_family="(T)n"
	repeat_region	complement(25045..25223)	/rpt_family="B4A"
	repeat_region	complement(25293..25414)	/rpt_family="RMER1B"
	repeat_region	25417..25585	/rpt_family="(GAA)n"
	repeat_region	26004..26237	/rpt_family="MLT1G1"
	repeat_region	26562..26722	/rpt_family="MLT1G"
	repeat_region	complement(27250..27583)	/rpt_family="TAPLTR1_MM"
	repeat_region	complement(27586..28359)	/rpt_family="TAPLTR1_MM-int"
	repeat_region	complement(28351..28707)	
Alignment Scores:			
Pred. No.:	2.1	Length:	198321
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0
US-09-858-332G-15 (1-12) x AC104863 (1-198321)			
Oy	1	HisaHnHisAenHiSAeNHISAeNHiSAeNHiSAeN	12
LOCUS	AC122454	198341 bp	DNA linear ROD 13-NOV-2003
DEFINITION	Mus musculus BAC clone RP24-263114 from chromosome 17, complete		
ACCESSION	AC122454		
VERSION	AC122454.4	GI:24476149	

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

HTG.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 198341)
Wang, C., Cotton, M., and Schatzkamer, K.
The sequence of Mus musculus BAC clone RP24-263114
2 (bases 1 to 198341)
Wilson, R.
Sequencing of Mus musculus
Unpublished (2001)
3 (bases 1 to 198341)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 198341)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 198341)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (02-NOV-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
6 (bases 1 to 198341)
Wilson, R.
Direct Submission
Submitted (13-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 2, 2002 this sequence version replaced gi:22476139.

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu

Summary Statistics
Center Project name: M_BB0263114

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RP24-24 BAC library has been constructed by Pieter de Jong and
coworkers (http://www.choi.org) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at http://www.choi.org

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.
Location/Qualifiers
1. 198341
/organism="Mus musculus"

/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="17"
/map="17"
/clone="RP24-263114"
/clone_lib="RP24-24"
11. .225
/rpt_family="Alu"
535. .642
repeat_region
552. .656
/rpt_family="B4"
repeat_region
859. .873
/rpt_family="B2"
874. .1020
repeat_region
1021. .1099
/rpt_family="Alu"
1153. .1327
repeat_region
1487. .1604
/rpt_family="B2"
repeat_region
1921. .2023
/rpt_family="MALR"
2120. .2259
repeat_region
2279. .2375
/rpt_family="Alu"
2415. .2635
repeat_region
2636. .2966
/rpt_family="MALR"
2967. .3027
repeat_region
3028. .3221
/rpt_family="MALR"
3310. .3358
repeat_region
3381. .3496
/rpt_family="B4"
3605. .3753
repeat_region
4396. .4422
/rpt_family="Alu"
4423. .4606
repeat_region
4607. .4770
/rpt_family="B2"
6602. .7046
repeat_region
7521. .7683
/rpt_family="ERVK"
8878. .8923
repeat_region
9046. .9101
/rpt_family="ID"
11754. .11883
repeat_region
11952. .12201
/rpt_family="RMR14"
12683. .12748
repeat_region
12756. .13101
/rpt_family="ID"
14642. .14771
repeat_region
14648. .14648
/rpt_family="Alu"
15058. .15248
repeat_region
15131. .15249
/rpt_family="MALR"
15250. .15296
repeat_region


```

repeat_region      /rpt_family="B4"
15439..15615
/rpt_family="B4"
repeat_region      /rpt_family="B4"
23433..23621
/rpt_family="B4"
repeat_region      /rpt_family="B4"
24052..24239
/rpt_family="B2"
repeat_region      /rpt_family="B2"
24261..24361
/rpt_family="Alu"
repeat_region      /rpt_family="Alu"
24344..24619
/rpt_family="B4"
repeat_region      /rpt_family="B4"
24641..24861
/rpt_family="B2"
repeat_region      /rpt_family="B2"
24876..25012
/rpt_family="L1"
repeat_region      /rpt_family="L1"
25180..25282
/rpt_family="Alu"
repeat_region      /rpt_family="Alu"
25737..25908
/rpt_family="L1"
repeat_region      /rpt_family="L1"
25779..25914
/rpt_family="L1"
repeat_region      /rpt_family="L1"
25918..26347
/rpt_family="L1"
repeat_region      /rpt_family="L1"
26638..26768
/rpt_family="Alu"
repeat_region      /rpt_family="Alu"
29289..29419
/rpt_family="Alu"
repeat_region      /rpt_family="Alu"
36950..37074
/rpt_family="MIR"
repeat_region      /rpt_family="B2"
37091..37281
/rpt_family="B2"
repeat_region      /rpt_family="B2"
44383..44475
/rpt_family="Alu"
repeat_region      /rpt_family="B2"
45853..46029
/rpt_family="B4"
repeat_region      /rpt_family="B4"
46938..47183
/rpt_family="MalR"
repeat_region      /rpt_family="B2"
47489..47522
/rpt_family="B2"
repeat_region      /rpt_family="B2"
52152..52291
/rpt_family="Alu"
repeat_region      /rpt_family="Alu"
53004..53079
/rpt_family="ID"
repeat_region      /rpt_family="ID"
53108..53205
/rpt_family="Alu"
repeat_region      /rpt_family="Alu"
53218..53280
/rpt_family="B2"
repeat_region      /rpt_family="B2"
53442..53413
/rpt_family="Alu"

```

Alignment Scores:

Pred. No.:	2	1	Length:	198341
Score:	84.00		Matches:	12
Percent Similarity:	100.00%		Conservative:	0
Best Local Similarity:	100.00%		Mismatches:	0
Query Match:	100.00%		Indels:	0
DB:	10		Gaps:	0

US-09-858-332G-15 (1-12) x AC122454 (1-198341)

QY 1 H1SASNH1SASNH1SASNH1SASNH1SASNH1SASN 12

DB 103983 CACAACCAACCAACCAACCAACCAACCAAC 104018

RESULT 70

LOCUS BX897729 201932 bp DNA linear VRT 27-JUL-2004

DEFINITION Zebrafish DNA sequence from clone DKEY-240A12 in linkage group 4.

ACCESSION BX897729

VERSION BX897729.6 GI:49614058

KEYWORDS HTG.

SOURCE Dario rerio (zebrafish)

ORGANISM Dario rerio

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Dario.

AUTHORS 1 (bases 1 to 201932)

TITLE Garnett, J.

JOURNAL Direct Submission

COMMENT Submitted (27-JUL-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Jul 1, 2004 this sequence version replaced gi:49457615.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the rare assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep

Clone-derived zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhifeng Bao and Sean Eddy, submitted), and those beginning 'dir' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-240A12 is from a zebrafish BAC library

VECTOR: pindigobac-5.

FEATURES

source

1..201932

/organism="Dario rerio"

/mol_type="genomic DNA"

/db_xref="taxon:7955"

/clone="DKEY-240A12"

/clone_lib="DantoKey"

ORIGIN

Alignment Scores:

Pred. No.:	2	14	Length:	201932
Score:	84.00		Matches:	12
Percent Similarity:	100.00%		Conservative:	0
Best Local Similarity:	100.00%		Mismatches:	0
Query Match:	100.00%		Indels:	0
DB:	5		Gaps:	0

US-09-858-332G-15 (1-12) x BX897729 (1-201932)

QY 1 H1SASNH1SASNH1SASNH1SASNH1SASNH1SASN 12

DB 112864 CACAATCAATCAATCAATCAATCAATCAAT 112899

RESULT 71

LOCUS AC069007/c 204940 bp DNA linear HTG 20-AUG-2000

DEFINITION	Homo sapiens chromosome 17 clone RP11-798G7, WORKING DRAFT
ACCESSION	SEQUENCE, 21 unordered pieces.
VERSION	AC069007.3 GI:9799889
KEYWORDS	HTS; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryotic; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 204940) Waterston,R.H.
TITLE	The sequence of Homo sapiens clone
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 204940)
AUTHORS	Waterston,R.H.
TITLE	Direct Submission
JOURNAL	Submitted (16-MAY-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63106, USA
COMMENT	On Aug 13, 2000 this sequence version replaced gi:3309540.

repeat_region	/rpt_family="L1ME1" complement(3902,.4200) /rpt_family="AluDb" complement(4201,.4279) /rpt_family="L1ME1" complement(4280,.4573) /rpt_family="AluY" complement(4576,.4881) /rpt_family="AluDb" complement(4882,.5105) /rpt_family="L1ME1" complement(5114,.5308) /rpt_family="MER58B" complement(5446,.5967) /rpt_family="L1MD4" complement(5974,.6104) /rpt_family="FLAM_C" complement(6114,.6415) /rpt_family="AluBx" complement(6439,.7049) /rpt_family="L1ME1" 7050,.7350 /rpt_family="AluBx" 7359,.7654 /rpt_family="AluBx" complement(7655,.7672) /rpt_family="L1ME1" complement(7797,.7913) /rpt_family="FLAM_C" complement(7951,.8309) /rpt_family="L1ME1" 8310,.8338 /rpt_family="TG)n" 8339,.8486 /rpt_family="TA)n" complement(8664,.8948) /rpt_family="AluSc" 8952,.9026 /rpt_family="L2" 9245,.9599 /rpt_family="L1MDA" 9623,.9738 /rpt_family="MIR" 9739,.10031 /rpt_family="AluBx" complement(10406,.10447) /rpt_family="L2" complement(11055,.11188 /rpt_family="MER5A" 11528,.11592 /rpt_family="MER5A" 11593,.11748 /rpt_family="AluSg" complement(11761,.11875) /rpt_family="FLAM_A" complement(11888,.12171) /rpt_family="AluBx" complement(13010,.13302) 13423,.13443 /rpt_family="AluSg" 13633,.13929 /rpt_family="AluY" complement(14020,.14174) /rpt_family="L2" 15545,.15618 /rpt_family="TTAA)n" complement(15619,.15901) /rpt_family="AluSc"
---------------	--

Alignment Scores:
Pred. No.:
Score:

2.21
84.00

Length:
Matches:

207611
12

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

```

QY      1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
      |||||
Db      172803 CATAATCATAATCATAATCATAATCATAAT 172768

```

RESULT 73	
AC120797/c	
LOCUS	AC120797
DEFINITION	Mus musculus chromosome 18, clone RP24-8D18, complete sequence.
ACCESSION	AC120797
VERSION	AC120797.8 GI:45544747
KEYWORDS	HTG.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (baees 1 to 210632)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus chromosome 18, clone RP24-81D18
Unpublished
2 (baees 1 to 210632)
Birren, B., Linton, L., Nusbaum, C., Lander, E.

AUTHORS
Bairren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastian, V., Bloom, T., Boguslavsky, I.,
Bouthalgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, T.,
Chazarab, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Deatrellano, K., Dowar, K., Diaz, J. S., Dodge, S.,
Farro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gordon, S., Goyette, M., Graham, L., Grand, Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamut, A., Karatas, A., Kells, C., Lacroque, K., Lamaszates, R.,
Landers, T., Lehoczyk, U., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, T., McEwan, P., McKernan, K., Meldrum, J., Menzies, L.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, T., Nguyen, C., Nicol, R.,
O'Connell, P., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunhphang, P., Pieterse, N., Pollara, P.,
Raymond, C., Retter, R., Ribick, W., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosettli, M., Roy, A., Santos, R., Schnauer, S., Schnapack, R.,
Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testa, J. S., Theodore, J.,
Topham, K., Travers, M., Travers, N., Trisigilo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. Y., Young, G.,
Zainoun, J., Zemek, L., Zimmer, A. and Zody, M.

JOURNAL Submitted (03-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 3 (bases 1 to 210632)

AUTHORS Birren, B., Nusbaum, C., Zander, E., Abouelela, A., et al.

Boguslavsky, L., Bounghalter, B., Camarata, J., Chiang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corman, B., DeCarliano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gadyra, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Houme, M., Iliev, I., Johnson, R., Jones, C., Kemat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Ton, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C., McDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mieng, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Ruchpala, A., Ramesamy, U., Raymond, C., Retts, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuppach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talmes, J., Teetaye, S., Theodore, S., Totham, K., Travers, M., Vassiliev, H., Venkatakrishnan, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zaitoun, J., Zemek, L., Zimmer, A. and Zody, M.

TITLE	Direct Submission
JOURNAL	Submitted (12-FEB-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	4 (bases 1 to 210632)
AUTHORS	Birren,B., Nusbaum,C., Lander,E., Abouelella,A., Allen,N., Anderson,M., Atchchil,H.M., Batra,N., Baetjen,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeAvellaon,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galgan,J., Garcia,S., Graham,L., Grand-Pierre,N., Hatzel,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrum,J., Meness,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasaamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Toppan,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,W.
TITLE	Direct Submission
JOURNAL	Submitted (18-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Mar 18, 2004 this sequence version replaced gi:42539078. All repeats were identified using RepeatMasker: Smit, A.P.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RV/RepeatMasker.html
TITLE	Genome Center
JOURNAL	Center: Whitehead Institute/ MIT Center for Genome Research
COMMENT	Center code: MIBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu Project Information Center project name: L24984 Center clone name: 81_D_18
FEATURES	----- Location/Qualifiers 1..210632 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" /chromosome="18" /map="18" /clone="RP24-81D18" /clone_1lb="RPCT-24 Male Mouse BAC" 1..8167 /note="wgs_end_extension clone_end=SP6" 20..106 /rpt_family="L1" complement(39..43) /note="<30 qual SNCL region" 107..129 /rpt_family="TTAA)n" 130..1207 /rpt_family="L1" 1206..1257 /rpt_family="L1MB8" 1264..1365 /rpt_family="GGGA)n" 1284..1336 /note="single clone coverage" 1517..1734 /rpt_family="L1MB8" complement(1793..2164) /rpt_family="MTD" 2472..2650 /rpt_family="B2_Mm2" 2651..2842

repeat_region	/rpt_family="B2_Mm3"	3160_..3304	/rpt_family="B2_Mm3"	3160_..3304
repeat_region	/rpt_family="TAAA)n"	complement(4192_..4438)	/rpt_family="TAAA)n"	complement(4192_..4438)
repeat_region	/rpt_family="MER102B"	complement(5055_..5240)	/rpt_family="MER102B"	complement(5055_..5240)
repeat_region	/rpt_family="B4"	5710_..5772	/rpt_family="B4"	5710_..5772
repeat_region	/rpt_family="(TCC)n"	5878_..6024	/rpt_family="(TCC)n"	5878_..6024
repeat_region	/rpt_family="B1_MM"	complement(6159_..6245)	/rpt_family="B1_MM"	complement(6159_..6245)
repeat_region	/rpt_family="MERS8B"	7018_..7060	/rpt_family="MERS8B"	7018_..7060
repeat_region	/rpt_family="(TG)n"	7060_..7094	/rpt_family="(TG)n"	7060_..7094
repeat_region	/rpt_family="(GA)n"	complement(7107_..7208)	/rpt_family="(GA)n"	complement(7107_..7208)
repeat_region	/rpt_family="ID_B1"	7271_..7468	/rpt_family="ID_B1"	7271_..7468
repeat_region	/rpt_family="B3"	complement(7500_..7628)	/rpt_family="B3"	complement(7500_..7628)
repeat_region	/rpt_family="BLACKXACK"	7650_..7699	/rpt_family="BLACKXACK"	7650_..7699
repeat_region	/rpt_family="(TTG)n"	complement(7704_..7808)	/rpt_family="(TTG)n"	complement(7704_..7808)
repeat_region	/rpt_family="ID_B1"	8011_..8114	/rpt_family="ID_B1"	8011_..8114
repeat_region	/rpt_family="(GTTG)n"	8168_..8171	/rpt_family="(GTTG)n"	8168_..8171
misc_feature	/note="clone_boundary clone_end:SP6 site:hdor"	complement(8657_..8730)	/note="clone_boundary clone_end:SP6 site:hdor"	complement(8657_..8730)
repeat_region	/rpt_family="L1ME"	complement(8876_..9355)	/rpt_family="L1ME"	complement(8876_..9355)
repeat_region	/rpt_family="BLACKXACK"	10244_..10524	/rpt_family="BLACKXACK"	10244_..10524
unsure	/rpt_family="L1_MM"	10426_..10505	/rpt_family="L1_MM"	10426_..10505
repeat_region	/note="single clone coverage"	10605_..10806	/note="single clone coverage"	10605_..10806
repeat_region	/rpt_family="B3"	11680_..11713	/rpt_family="B3"	11680_..11713
repeat_region	/rpt_family="(TAAA)n"	12090_..12194	/rpt_family="(TAAA)n"	12090_..12194
repeat_region	/rpt_family="L12"	12396_..12435	/rpt_family="L12"	12396_..12435
repeat_region	/rpt_family="(TTTG)n"	complement(13067_..13387)	/rpt_family="(TTTG)n"	complement(13067_..13387)
repeat_region	/rpt_family="ORR1B1"	13679_..13707	/rpt_family="ORR1B1"	13679_..13707
repeat_region	/rpt_family="(CA)n"	complement(13718_..13769)	/rpt_family="(CA)n"	complement(13718_..13769)
repeat_region	/rpt_family="ORR1A1"	13783_..13811	/rpt_family="ORR1A1"	13783_..13811
repeat_region	/rpt_family="AT_rich"	complement(13815_..15162)	/rpt_family="AT_rich"	complement(13815_..15162)
repeat_region	/rpt_family="L1_MM"	complement(15163_..15448)	/rpt_family="L1_MM"	complement(15163_..15448)
repeat_region	/rpt_family="ORR1A2"	complement(15449_..15532)	/rpt_family="ORR1A2"	complement(15449_..15532)
repeat_region	/rpt_family="L1MC4"	complement(15504_..15579)	/rpt_family="L1MC4"	complement(15504_..15579)
repeat_region	/rpt_family="L1MC4"	complement(15580_..15604)	/rpt_family="L1MC4"	complement(15580_..15604)
repeat_region	/rpt_family="B2L_S"		/rpt_family="B2L_S"	
Alignment Scores:				
Pred. No.:	2.24	Length:	2106322	
Score:	84.00	Matches:	12	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	

*	255.07	282.91	contig of 2785 bp in length
*	282.92	283.91	gap of unknown length
*	283.92	316.05	contig of 3214 bp in length
*	316.06	317.05	gap of unknown length
*	317.06	361.10	contig of 4405 bp in length
*	361.11	362.10	gap of unknown length
*	362.11	387.73	contig of 2563 bp in length
*	387.74	388.73	gap of unknown length
*	388.74	439.27	contig of 5054 bp in length
*	439.28	440.27	gap of unknown length
*	440.28	480.86	contig of 4059 bp in length
*	480.87	481.86	gap of unknown length
*	481.87	520.42	contig of 3856 bp in length
*	520.43	521.42	gap of unknown length
*	521.43	562.65	contig of 4123 bp in length
*	562.66	563.65	gap of unknown length
*	563.66	614.49	contig of 5084 bp in length
*	614.50	615.49	gap of unknown length
*	615.50	723.66	contig of 10819 bp in length
*	723.69	724.66	gap of unknown length
*	724.69	793.29	contig of 6661 bp in length
*	793.30	794.29	gap of unknown length
*	794.30	851.12	contig of 5683 bp in length
*	851.13	852.12	gap of unknown length
*	852.13	919.71	contig of 6759 bp in length
*	919.72	920.71	gap of unknown length
*	920.72	1026.75	contig of 10604 bp in length
*	1026.75	1027.75	gap of unknown length
*	1027.76	1125.43	contig of 9768 bp in length
*	1125.44	1126.43	gap of unknown length
*	1126.44	1305.22	contig of 17881 bp in length
*	1305.25	1306.22	gap of unknown length
*	1306.25	1436.80	contig of 13056 bp in length
*	1436.81	1437.80	gap of unknown length
*	1437.81	1692.27	contig of 25447 bp in length
*	1692.28	1693.27	gap of unknown length
*	1693.28	1933.40	contig of 24113 bp in length
*	1933.41	1935.40	gap of unknown length
*	1935.41	2200.06	contig of 26466 bp in length

FEATURES	Location/Qualifiers
source	1. .220006

11

```
/clone_lib="RPC1 mouse BAC library 23"
```

Alignment Scores:

Pred. No.:	2,34	Length:	220006
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-858-332G-15 (1-12) X AC084068 (1-220006)

QY 1 HisAenHisAsnHisASnHisASnHisASn 12
 |||
Db 47332 CATAACCATAACCATTAACCATTAACCATTAAC 47367

RESULT	76		
AC073680/c			
LOCUS	AC073680	226013 bp	DNA
DEFINITION	Mus musculus clone RP23-116M1, WORKING DRAFT SEQUENCE. 20 unordered	linear	HTG 29-JUN-2000

ACCESSION	AC073680
VERSION	AC073680.1
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
1	Bhakarjora; Metazoa;	Chordata; Craniata;	Eureleostomi;	1	(bases 1 to 226013)		
2	Mammalia; Eutheria;	Rodentia; Sciurognathi;	Muridae; Murinae; Mus	2	(bases 1 to 226013)		
	Sequencing of Mouse				Unpublished		
	DOE Joint Genome Institute.				DOE Joint Genome Institute.		
	Submitted (29-JUN-2000)	Production Sequencing Facility,	DOE Joint				
	Genome Institute, 2800 Mitchell Drive,	Walnut Creek, CA 94598, USA.					
	-----Genome Center						

```
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 1781705
Center clone name: RPCI-23_116M1
```

Center Project Name: 1781705
Center clone name: RPCI-23_116M1

Summary Statistics

Consensus quality: 209118 bases at least Q40

Consensus quality: 215883 bases at least Q30

Consensus quality: 217616 bases at least Q20

Estimated insert size: 216000; agarose-fp estimation

Estimated insert size: 224113; sum-of-contigs estimation

Quality coverage: 9.65 in Q20 bases; agarose-fp estimation

Quality coverage: 9.3 in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently

* consists of 20 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1	1157: contig of 1157 bp in length
1158	1257: gap of unknown length
1258	2308: contig of 1051 bp in length
2309	2408: gap of unknown length
2409	3408: contig of 1000 bp in length
3409	3508: gap of unknown length
3509	4765: contig of 1257 bp in length
4766	4865: gap of unknown length
4866	6182: contig of 1317 bp in length
6183	6282: gap of unknown length
6283	7697: contig of 1415 bp in length
7698	7797: gap of unknown length
7798	9195: contig of 1398 bp in length
9196	9295: gap of unknown length
9296	10579: contig of 1284 bp in length
10580	10679: gap of unknown length
10680	12073: contig of 1394 bp in length
12074	12173: gap of unknown length
12174	13684: contig of 1511 bp in length
13685	13784: gap of unknown length
13785	14932: contig of 1148 bp in length
14933	15032: gap of unknown length
15033	17640: contig of 2608 bp in length
17641	17740: gap of unknown length
17741	19342: contig of 1602 bp in length
19343	19442: gap of unknown length
19443	21366: contig of 1924 bp in length
21367	21465: gap of unknown length
21467	22246: contig of 5780 bp in length
22247	27346: gap of unknown length
27347	30988: contig of 3642 bp in length
30989	31088: gap of unknown length
31089	35750: contig of 4662 bp in length
35751	35850: gap of unknown length
35851	50689: contig of 14839 bp in length
50689	50789: gap of unknown length
50790	70278: contig of 19485 bp in length

```
repeat_region 2476.  .2564
```


	repeat_region	complement(119925..12030)	/rpt_family="B1P"	
	repeat_region	complement(12265..12445)	/rpt_family="B3"	
	repeat_region	complement(12536..12731)	/rpt_family="B3"	
	unSURE	complement(12878..12880)	/note="single clone coverage"	
 Alignment Scores:				
Pred. No.:	2.42	Length:	227081	
Score:	84.00	Matches:	12	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	100.00%	Indels:	0	
DB:	10	Gaps:	0	
 US-09-8558-332G-15 (1-12) x AC119847 (1-227081)				
Oy	1	Hlsasnhlsasnhlsasnhlsasnhlsasnhlsasn	12	
Db	192501	CATACCATTAACCATTAACCATTAACCATTAAC	192466	
RESULT 78				
AC133743				
LOCUS				
DEFINITION	AC133743	227648 bp	DNA linear HTG 20-NOV-2002	
ACCESSION	AC133743			
VERSION	AC133743.2	GI:25139560		
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.			
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM	Rattus norvegicus			
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;			
	Rattus.			
REFERENCE	1 (bases 1 to 227648)			
AUTHORS	Mzyny,D,Marle, Metzker,M,Lee, Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Albrooke,S., Amin,A., Angulo,D., Ayalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,K., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biewal,K., Blair,J., Blankenburg,K., Blych,P., Brown,M., Bryant,N., Bunay,C., Burck,P., Burrell,K., Calderon,B., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Devila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Evans,K., Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Gante,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,R., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Haylak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,B., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Huily,S., Hume,J., Idlebird,D., Jackson,A., Jackson,B., Jacob,L., Jiang,H., Johnson,B., Johnson,D., Jollivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lepow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,T., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuwa,L., Louissegh,H., Lozada,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartine,M., Mahmoud,M., Mallory,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,B., Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenan,E., Milosavljevic,A., Miner,G., Minia,E., Montemeyor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Mwinda,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaekeleneh,O., Okunonu,G., Olampongsooi,A., Pal,S., Parks,K., Patelena,S., Paul,H., Perez,A., Perez,L., Phannoch,C., Plopper,F., Polndexter,A., Popovic,D., Primus,E., Pu,L.-L., Piazzi,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Relliy,B., Relliy,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,			

[illegible]

Submitted: 12 November 2002; Human Genetics, Baylor College of Medicine, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 20, 2002 this sequence version replaced gi:23096694.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/Projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

FEATURES	SOURCE
*	220175
*	220274: gap of unknown length
*	220275
*	221473: contig of 1159 bp in length
*	221474
*	221573: gap of unknown length
*	221574
*	223286: contig of 1713 bp in length
*	223287
*	223386: gap of unknown length
*	224587: contig of 1201 bp in length
*	224588
*	224687: gap of unknown length
*	225923: contig of 1236 bp in length
*	225924
*	227648: contig of 1625 bp in length
*	226024
	location/Qualifiers
	1..227648
	/organism="Rattus norvegicus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10116"
	/clone="CH230-182C23"
	129007..130011
	/note="wgs contig"
	155940..167729
	/note="wgs contig"
	198830..199863
	/note="wgs contig"
	203077..204817
	/note="wgs contig"

Alignment Scores:

Alignment Scores:	
Pred. No.:	2.43
Score:	84.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	2
Length:	227648
Matches:	12
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-858-332G-15 (1-12) X AC133743 (1-227648)

1 HIBASHNH1BASNH1BASNH1BASNH1BASn 12

[illegible]

AC096973/

DEFINITION Rattus norvegicus clone CH230-2N7, WORKING DRAFT SEQUENCE, 3
unorderd nt

ACCESSION
VERSION

SOURCE	Rattus norvegicus (Norway rat)
	HIGS_FULLTOP; HIGS_DRAFT; HIGS_FINAL

Rattus.
Rattus.

AUTHORS

1 (passes to 23183)

Muzny, D., Marie, Metzger, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnesstead, M., Benahmed, F., Benwal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesaar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dedrick, D., Delgado, O., Benson, S., Desarmo, C., Ding, Y., Dinh, R., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Dublin, K., Duval, B., Evans, K., Egan, A., Escott, M., Eugene, C., Evans, C., Fallett, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, P., Greengrass, E., Geer, K., Giller, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Halland, W., Hamill, C., Hamilton, C., Hamilton, C., Harrey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,

Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebirt, D., Jackson, A., Jackson, S., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kover, C., Kowis, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, D., Lorenz, L., Louie, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., McNeill, T., Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okunolu, G., Olajunwa, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Piopier, F., Polidexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quirroz, J., Rachin, E., Reeves, K., Regier, M., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojars, A., Rose, M., Rose, R., Ruiz, S., J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C., D., Smajic, D., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D., R., Holt, R., A., Smith, H., O., Weinstein, G., and Gibbs, R., A.

TITLE

Unpublished
Direct Submission

2 (bases 1 to 231383)

Worley, K.C.

Direct Submission

Submitted (05-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 231383)

Baylor College of Medicine

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 13, 2002 this sequence version replaced gi:23683156.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas

(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the end of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GAWT

Center clone name: CH230-2N7

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 216032 bases at least Q40

Consensus quality: 21638 bases at least Q30

Consensus quality: 218825 bases at least Q20

Estimated insert size: 223315; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 215694: contig of 215694 bp in length
* 2 215695 215794: gap of unknown length
* 3 215795 218857: contig of 3063 bp in length
* 4 218858 218957: gap of unknown length
* 5 218958 231383: contig of 12426 bp in length.
* Location/Qualifiers
* 1. 231383
* /organism="Rattus norvegicus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-2N7"
* 1. 1209
* /note="wgs contig"
* complement(206387..206798)
* /note="clone boundary
* clone_end:Sp6
* site:
* end_sequence:BH289712"
* 211671..215694
* /note="wgs_end_extension
* clone_end:Sp6"

FEATURES

Source

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

Egan, A., Escotto, M., Eugene, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabiet, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guertler, M., Guevara, W.,
Gunnarctate, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, J., Havlik, P., Haves, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, W.,
Hollins, B., Howells, S., Huik, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Krat, C.L., Lebow, H., Levay, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lounsbuher, L., Louised, H., Lozada, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A.,
Mangum, B., Mapa, P., Martin, K., Martin, R., Martineau, E.,
Manwiny, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Mnjaja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Mundaasa, M., Murphy, M., Nair, L.,
Nankervyis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwachukwem, O., Okwoum, G., Olarnpusagoun, A., Pal, S., Parks, K.,
Pasaternak, S., Paul, H., Perez, A., Perez, L., Piamkoon, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Reiser, M.A., Reigh, R.,
Kelly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,
Sanders, W., Saverly, G., Scherer, S., Scott, G., Shateman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Siltner, C.D., Smag, D.,
Sneedy, A., Sodergren, E., Song, X.-Z., Soreiller, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Syntek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Veta, V., Villanasa, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausen, A., Weiser, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Direct Submission

TITLE Direct Submission
 JOURNAL
 REFERENCE 2 (bases 1 to 232413)
 AUTHORS Morley, K.C.
 TITLE Direct Submission
 JOURNAL Submitted (08-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 232413)
 AUTHORS Rat Genome Sequencing Consortium.
 TITLE Direct Submission
 JOURNAL Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

of Molecular and Human Genetics, Baylor Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23194579.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Aiala (<http://www.hgsa.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Aiala assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: GIKW
 Center clone name: CH230-268L9
 Summary Statistics

```

Assembly program: Phrap; version 0.990329
Consensus quality: 215301 bases at least Q40
Consensus quality: 214075 bases at least Q30
Consensus quality: 215384 bases at least Q20
Estimated insert size: 215514; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.bsc.bcm.tmc.edu/docs/genbank\_draft\_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

FEATURES	
*	230071: contig of 2343 bp in length.
*	229971: gap of unknown length
*	229970: contig of 1916 bp in length
*	228055: gap of unknown length
*	227954: contig of 5201 bp in length
*	227754: contig of 22763 bp in length
*	199962: gap of unknown length
*	199852: contig of 22702 bp in length
*	199851: gap of unknown length
*	189235: contig of 10617 bp in length
*	189234: gap of unknown length
*	42662: contig of 146443 bp in length
*	42591: gap of unknown length
*	42592: contig of 42591 bp in length

```

source
1. 232413
   /organism="Rattus norvegicus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10116"
   /clone="CH230-268L9"

```

```
misc_feature 1. .1737
              /note="wgs end_extension
clone end: T7"
misc_feature complement (4465. .5309)
              /note="clone_boundary
clone end: T7
```

```

misc_feature      end_sequence:BZ173680"
                  189235. 191036

```

```
misc_feature /note="wgs contig"
misc_feature 197729. .198806
misc_feature /note="wgs contig"
misc_feature 199952. .201761
```

```
misc_feature    /note="wgs contig"
                222754. .22182
                /note="wgs contig"
```

```

424233.421954
/note="wgs_contig"
N

```

Alignment Scores:	
Pred. No.:	2, 48
Score:	84.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	2
Length:	232413
Matches:	12
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-858-332G-15 (1-12) x AC117096 (1-232413)

vy 1 HISASNHISASHISASHISASHISASHNISHBN 12
 |||||
 Db 228695 CACAACCAACCAACCAACCAACCAACCAAC 228730
 |||||

RESULT 81	
AC127135/c	
DLOCUS	
AC127135	
237732 bp	DNA
14-0000	

DEFINITION Rattus norvegicus clone CH230-89C20. *** SEQUENCING IN PROGRESS
 *** 3 unordered pieces.
 ACCESSION AC127135
 VERSION AC127135.4 GI:30520824
 KEYWORDS HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 237732)
 Wuzny,D.,Marle, Metzker,M.,Lee, Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooke,S., Amin,A., Angiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Bismato,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhey,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Diya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Giller,R., Grady,M., Guerra,M., Guevara,W., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M., Hollins,B., Howells,S., Huily,H., Hume,J., Idlebirt,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,D., Lorensenwa,L., Loulsegged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mahoney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Nair,L., Nankarva,S.C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaelelehen,O., Okwunnu,G., Olarnpungsoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Polindexter,A., Popovic,D., Primus,E., Pu,L.-L., Pizzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richard,S., Riggs,F., Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,M., Saverly,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sison,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejo,J., Usmant,K., Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczky,R., Woodson,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausen,A., Weis,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstein,G., and Gibbs,R.A.

REFERENCE 2 (bases 1 to 237732)
 Unpublished
 Direct Submission
 2 (bases 1 to 237732)
 Morley,K.C.
 Direct Submission
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 237732)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One

COMMENT

Baylor Plaza, Houston, TX 77030, USA
 On May 10, 2003 this sequence version replaced GI:24819332. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rac/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GHZP
 Center clone name: CH230-89C20
 ----- Summary Statistics
 Assembly program: Atlas 3.0?
 Consensus quality: 206671 bases at least Q40
 Consensus quality: 211484 bases at least Q30
 Consensus quality: 214331 bases at least Q20
 Estimated insert size: 220281; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently

* consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 234751: contig of 234751 bp in length
 * 234752 234851: gap of unknown length
 * 234852 235859: contig of 1008 bp in length
 * 235860 235959: gap of unknown length
 * 235960 237732: contig of 1773 bp in length.

FEATURES

Location/Qualifiers
 1. 237732
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"

misc_feature

1. 1165
 /note="wgs_end-extension
 clone_end:T7"

misc_feature

complement(7252..8117)
 /note="clone boundary
 clone_end:T7"

misc_feature

end_sequence:BH337952"
 135628..136318
 /note="wgs contig"

misc_feature

164949..167852
 /note="wgs contig"
 229390..230234
 /note="clone boundary
 clone_end:Sp6"

misc_feature

end_sequence:BH337953"
 230944..234751
 /note="wgs_end-extension
 clone_end:Sp6"

ORIGIN

Pred. No.:	2	Length:	23773
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Genes:	0

US-09-858-3326-15 (1-12) X AC127135 (1-237732)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
 Dd |||||
 198279 CACAACCAACCAACCAACCAACCAACCAAC 198244

RESULT 8:
CR450845

LOCUS		23818 bp	DNA	linear	HTG 30-MAY-2007
DEFINITION	Danio rerio clone DKEY-65K7, *** SEQUENCING IN PROGRESS ***, 10				
UNORDERED PIECES.					
CPAE084c					
ACCESSION					

VERSION	CR450845.3	GI:47825483
KEYWORDS	HTG; HTGS PHASE1.	
SOURCE	Danio rerio (zebrafish)	
ORGANISM	Danio rerio	

REFERENCE
1 (bases 1 to 238318)
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE
AUTHORS
TITLE
JOURNAL

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 zfish_help@anger.ac.uk
 Clone requests: clonerequest@anger.ac.uk
 On May 30, 2004 this sequence version replaced gi:47775883.
 ----- Genome Center -----

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK63K7

```

----- Summary Statistics -----
Assembly program: XGAP4, version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 235122 bases at least Q40
Consensus quality: 235609 bases at least Q30
Consensus quality: 236075 bases at least Q20
Insert size: 237418; sum-of-contigs
Insert size: 233842; 3.0% error; agarose-fp
Quality coverage: 8.69% in Q20 bases; sum-of-contigs
Quality coverage: 8.83% in Q20 bases; agarose-fp

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

*		65303:	contig of 65303 bp	in length
*	1	65304:	gap of 100 bp	
*	65404	93550:	contig of 28147 bp	in length
*	93551	93650:	gap of 100 bp	
*	93651	108211:	contig of 14561 bp	in length
*	108212	108311:	gap of 100 bp	
*	108312	144758:	contig of 35487 bp	in length
*	144759	144898:	gap of 100 bp	
*	144899	156445:	contig of 11547 bp	in length
*	156446	156545:	gap of 100 bp	
*	156546	190876:	contig of 34331 bp	in length
*	190877	190976:	gap of 100 bp	
*	190977	207444:	contig of 16466 bp	in length

```

* 207544: gap of 100 bp
* 207545: 207544: confg of 3100 bp in length
* 210645: 210644: gap of 100 bp
* 210743: 210744: confg of 20299 bp in length
* 231043: 231044: gap of 100 bp
* 231044: 231043: confg of 7175 bp in length..
* 238118: Location/Qualifiers
  1: 238318

```

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

m3bc_feature

14-00000

ATMOSPHERIC

ORIGIN

Alignment Scores:

```

Pred. NO.:      2.55      Length:      288318
Score:          84.00      Matches:      12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00% Indels: 0
DB:             2        Gaps: 0

```

US-09-858-332G-15 (1-12) X CR450845 (1-238318)

[illegible]

RESULT 83
AC097541

DEF IN 110

KEYWORDS

100

REFERENCE

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 240663)

AUTHORS
 Muzny, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Arguiano, D., Ayalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, D., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Devita, M., L., Davis, C., Davy-Carroll, L., De Anda, C., Dedrich, D., Delgado, O., Denon, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haland, M., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hui, S., Hume, J., Idubird, D., Jackson, A., Hollins, B., Howell, S., Hui, S., Hume, J., Idubird, D., Jolivet, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kary, C., Karpach, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, D., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, L., Lorensuwa, L., Louised, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Mirza, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Naochalemech, O., Okunolu, G., Olarnunagong, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Piopier, F., Polidexter, A., Popovic, D., Pimm, E., Pu, L., L., Puzo, M., Quiriz, J., Rachin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatk, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wleczek, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Weinstock, G., and Gibbs, R. A.

TITLE
 Unpublished

JOURNAL
 2 (bases 1 to 240663)

REFERENCE
 Worley, K.C.

AUTHORS
 Direct Submission

TITLE
 Submitted (19-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

JOURNAL
 3 (bases 1 to 240663)

REFERENCE
 Rat Genome Sequencing Consortium.

AUTHORS
 Direct Submission

TITLE
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

JOURNAL
 On May 10, 2003 this sequence version replaced gi:2495666.

COMMENT
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information
 Center project name: GGGT
 Center clone name: CH230-2701

----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 215135 bases at least Q40
 Consensus quality: 220155 bases at least Q30
 Consensus quality: 223518 bases at least Q20
 Estimated insert size: 228408; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 232805: contig of 232805 bp in length
 * 232806 232905: gap of unknown length
 * 232906 234010: contig of 1105 bp in length
 * 234011 234110: gap of unknown length
 * 234111 235442: contig of 1432 bp in length
 * 235443 235642: gap of unknown length
 * 235643 237928: contig of 2286 bp in length
 * 237929 238028: gap of unknown length
 * 238029 240663: contig of 2635 bp in length.

FEATURES
 location/Qualifiers

1..240663
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-2701"

1..2549
 /note="wgs_end_extension
 clone_end:Sp6
 site:ECORI
 end_sequence: BH337345"
 4496..74792
 /note="clone_boundary
 clone_end:T7
 site:BCORI
 end_sequence: BH337344"
 60177..61683
 /note="wgs_end_extension
 clone_end:T7"
 108980..111715
 /note="wgs_end_extension
 clone_end:T7"
 227776..232805
 /note="wgs_end_extension
 clone_end:T7"

ORIGIN
 Alignment Scores:
 Pred. No.: 2.57 Length: 240663
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

repeat_region complement (35720..35788)
 Alignment Scores:
 Pred. No.: 2.61 Length: 243772
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-858-332G-15 (1-12) x AY491413 (1-2433772)

Qy	1	HISaenhISAenhISAenhISAenhISAenhISA	12
DB	243636	CACAAACCAACCAACCAACCAACCAACCAACCAAC	243671
RESULT 86			
AC094502/c			
LOCUS			
DEFINITION	AC094502	248752 bp	DNA linear HTG 09-MAY-2003
ACCESSION	AC094502		
VERSION	AC094502.6		
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.		
SOURCE	Rattus norvegicus clone CH230-4E8, *** SEQUENCING IN PROGRESS ***		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 248752)		
AUTHORS	Muzny, D., Marie, M., Metzker, M., Lee, A., Adamson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angilano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Bace, E., Baden, H., Baidya, B., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Gatta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K., Harey, J., Hawik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., Hollins, B., Howell, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpaty, S., Kelly, S., Kelly, S., Khan, Z., King, L., Koyar, C., Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, Y., Liu, Y., London, P., Longacre, S., Lopez, U., Lorenshewa, L., Louissed, H., Lozard, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Matisoff, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Ming, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nowoklenneth, O., Okwunonu, G., Olamunnsagoon, A., Pal, S., Parks, K., Patenak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L., Pu, Z., Puzo, M., Quito, J., Rachlin, E., Reeves, K., Regier, M. A., Reich, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, E., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. U., Sanders, W., Savary, G., Scherer, S., Scott, G., Shattman, S., Shen, H., Shetty, J., Shvartberg, A., Sison, I., Sitter, C. D., Smaj, D., Sneed, A., Sodegren, E., Song, X.-Z., Sorelle, R., Soosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, K., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Treloar, Z., Umami, K.,		

Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Woodan, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G., and Gibbs, R.A.

TITLE
Journal
Unpublished
2 (bases 1 to 248752)

REFERENCE
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 248752)
Rat Genome Sequencing Consortium.

COMMENT
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24818812.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GATW
Center clone name: CH230-4E8

----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 206523 bases at least Q40
Consensus quality: 21412 bases at least Q30
Consensus quality: 214233 bases at least Q20
Estimated insert size: 215006; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 245823: contig of 245823 bp in length
245824 245923: gap of unknown length
245924 247299: contig of 1376 bp in length
247300 247399: gap of unknown length
247400 248752: contig of 1353 bp in length.

FEATURES
source
1. 248752
/mol_type="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-4E8"
1. 1697
/note="wgs_end_extension
clone_end:T7"

misc_feature
2836..3642
/note="clone_boundary
clone_end:T7
site:ECORI
end_sequence:BH305316"
misc_feature
90194..91759
/note="wgs contig"
91810..93031
/note="wgs contig"
152801..154271
/note="wgs contig"
184165..185672
/note="wgs contig"
complement(1237880..238728)
/note="clone_boundary
clone_end:Spe
site:ECORI
end_sequence:BH305317"
244828..245823
/note="wgs_end_extension
clone_end:Spe"

misc_feature
misc_feature
misc_feature
misc_feature
misc_feature

ORIGIN
Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
2
66
84.00
100.00%
100.00%
2

US-09-858-332g-15 (1-112) x AC094502 (1-248752)

QY
1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

Db
136235 CACCAACCAACCAACCAACCAACCAACCAAC 136200

RESULT
87
AC113885/c
LOCUS
DEFINITION
AC113885
Rattus norvegicus clone CH230-336824, *** SEQUENCING IN PROGRESS
***, 2 unordered pieces.
ACCESSION
AC113885.5 GI:24635526
VERSION
KEYWORDS
HTG, HTGS_PASSEL, HTGS_DRAFT, HTGS_ENRICHED.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 248784)
Wozny, D., Matre, V., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, B., Baden, H., Baldwin, D., Bardaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dedrich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Bayes, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Giller, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

DEFINITION Rattus norvegicus clone CH230-4G23, WORKING DRAFT SEQUENCE, 2 unordered pieces.

ACCESSION AC126742

VERSION AC126742.6 GI:30467824

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 253494)
Muzny,D,Marle, Metzker,M,Lee, Abramson,S, Adams,C, Alder,J, Allen,C, Allen,H, Alsbrooks,S, Amin,A, Angiano,D, Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,D, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benamed,F, Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Bunay,C, Burch,P, Burrell,K, Calderon,E, Cardenas,V, Carter,K, Cavazos,I, Ceasar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Diya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Gebregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W, Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,J, Jackson,B, Howell,S, Hulky,S, Hume,J, Idlebird,D, Jackson,A, Johnson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpach,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J, Liu,Y, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorenshuwa,L, Louisedge,H, Lozado,R,J, Lu,X, Ma,J, Maheshwari,M, Mahindartine,M, Mahmoud,M, Malloy,K, Mangum,A, Mangum,B, Mapua,P, Martin,K, Martin,R, Martineau,E, McWhiney,S, McLeod,M,P, McNeill,T,Z, Meenen,E, Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Ngunidasa,M, Murphy,M, Nair,L, Nankervis,C, Neal,D, Newton,N, Niyidasa,N, Norris,S, Nwaokemelehu,O, Okwunu,G, Olarunpasegun,A, Pal,S, Parks,K, Paternak,S, Paul,H, Perez,A, Perez,L, Pfannkuch,C, Plopper,F, Poindexter,A, Popovic,D, Primus,E, Pu,L-L, Pizzo,M, Quiroz,J, Rachin,E, Reeves,K, Regier,M,A, Reigh,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,W, Savery,G, Scherer,S, Scott,G, Shatsman,S, Shen,H, Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smajic,D, Speed,A, Sodergren,E, Song,X-Z, Sorelle,R, Sosa,J, Steinle,M, Strong,R, Sutton,A, Svatek,A, Tabor,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Thunge,A, Trejos,Z, Usmani,K, Valas,R, Vera,V, Villaseana,D, Waldron,L, Walker,B, Wang,D, Wang,O, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Willison,R, Wleczky,R, Wooden,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von Niederhausen,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstein,G, and Gibbs,R,A.

TITLE Unpublished

JOURNAL 2 (bases 1 to 253494)

REFERENCE Direct Submission

AUTHORS Worley,K,C.

JOURNAL Submitted (09-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 253494)

AUTHORS Rat Genome Sequencing Consortium.

TITLE Direct Submission

JOURNAL Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One

COMMENT

Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:24941150. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GCGW

Center clone name: CH230-4G23

----- Summary Statistics

Assembly program: Atlas;

Consensus quality: 212667 bases at least Q40

Consensus quality: 215613 bases at least Q30

Consensus quality: 217844 bases at least Q20

Estimated insert size: 224651; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preereved.

* 1 252279: contig of 252279 bp in length

* 252280 252379: gap of unknown length

* 252380 253494: contig of 1115 bp in length.

Location/Qualifiers

1..253494

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-4G23"

1..1098

/note="wgs contig"

misc_feature

2071..4080

/note="wgs contig"

4131..5739

/note="wgs contig"

ORIGIN

Alignment Scores:

Pred. No.: 2.72

Score: 84.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 100.00%

DB: 2

US-09-858-332G-15 (1-12) x AC126742 (1-253494)

Oy 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

|||||

Db 204196 CACACCAACACCAACCAACCAACCAACCAAC 204231

|||||

RESULT 91

AC117075

[illegible]

CDS

DWEASSKDRINVDNVFFSVIKRIKRYKKEGAPVYKKEKCIIM"
12794. .13423
/note="Geneid exon scores (in order of location ranges):
35.19 - GSCJ_ID dd_00410"

CDS

/codon_start=1
/product="hypothetical protein"
/protein_id="AA050736.1"
/db_xref="GI:28828052"
/translation="MASIGTKFKNEPTFGKNIKELILERSIAKSEKOWDLVSEN
KNLVNAPRINKLFDNKIGLDYDKNSPEFLIKTYHEPTEPVIDIYNKSCIPY
LSDLQSSITITNPENIQAOKDSTVPFARLNSISKNVQIKRHYLSNRITELVC
CLILPANIDITCKVVPKSEFLNETSISGCPFLMDSSPYGKKLST"
complement (join(13494. .13813,13915. .14097,14208. .14241))
/note="Geneid exon scores (in order of location ranges):
22.37, 12.24, 8.90 - GSCJ_ID dd_00408"

CDS

/codon_start=1
/product="hypothetical protein"
/protein_id="AA050737.1"
/db_xref="GI:28828053"
/translation="MSIYEIRMAOVKOGMGSGPLGYILRENNKKEKAVKAVNP
KPKPVVOTSPYITINIEGTVTATDNNGSDPYLITAIKPNNSFKILFKSNCKK
TLTPMNQVENPFVKDNDIYELMDHVDLSKNDIIGSIISMNTVRSKGYTIEF
FMSEVVDGIVKLTIRK"
complement (join(14655. .14686,14848. .14916,14992. .15162,
15268. .15310))
/note="Geneid exon scores (in order of location ranges):
1.29, 3.37, 1.88, 4.01 - GSCJ_ID dd_00406"

CDS

/codon_start=1
/product="hypothetical protein"
/protein_id="AA050738.1"
/db_xref="GI:28828054"
/translation="MSEICSTSCMOYAIIDNPNCPIGTFKVCFMGSNTYTCASSGFT
VTSDELSPINVGHYHDEILIGKRIKIKNGEYKTKATNLSINLSNKNRKYCLPBRK
V"
complement (join(15876. .15900,16018. .17390))
/note="Geneid exon scores (in order of location ranges):
-0.56, 167.37 - GSCJ_ID dd_00405"

CDS

/codon_start=1
/product="similar to Physarum polycephalum (Slime mold).
Glutathione reductase"
/protein_id="AA093954.2"
/db_xref="GI:28828055"
/translation="MSTNHFYLVLAGSGGIASARAKHLNKGNGRIGIVEVT
RGGTCVNVGCVKPMNNTSFIKEMINAPSYGDFGQGVFEMPTIKKADKDEYK
RLNGIYDLSLAKNDIYIRNGYRSGPKIEOVNANGKXTADHLLIAGGRTVDV
PKKEIGITSDGFELEDPKSTIVVAGIYAVELAVHSISGETTMRQKPLATF
DEMLATTLKQMTDGVKVTETASISLESDVDGKRIATNAGVLPVECVTAIG
RVPTNDLIGIDKAGIQLTQSGFIKYDEQNTMVPVHAGVDCGFLTPVIAIAGR
RLSERLFNKSGSDIKFEYENAVATVFSHPITGVLTEGEATITGTEINIKCVNTSFIN
MFYSVQVHKVRTSMKLVLCGKEKVKVIGLHIGDCDBIIGQFAVAVKMGCTKMDLNT
CAIHPTSAEELVTMV"
complement (join(18106. .18501,18638. .18994))
/note="Geneid exon scores (in order of location ranges):
56.32, 49.02 - GSCJ_ID dd_00404"

Alignment Scores:

Pred. No.: 2 73 length: 254733
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-858-332g-15 (1-12) x AC117075 (1-254733)

Qy 1 H18AsnH18AsnH18AsnH18AsnH18AsnH18Asn 12

Db 129205 CACAACCAACCAACCAACCAACCAACCAAC 129240

RESULT 92

AC107178 255032 bp DNA linear HTG 08-OCT-2002
LOCUS
DEFINITION Rattus norvegicus clone CH230-162F7, *** SEQUENCING IN PROGRESS

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISMREFERENCE
AUTHORS

***, 2 unordered pieces.
AC107178 GI:22856980
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus

1 (bases 1 to 255032)

Wozny,D,Marie, Metzker,M, Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benhammed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Bacotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregregoris, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gumarate, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M.,
Hollins, B., Howells, S., Hui, S., Hume, J., Idlebird, D., Jackson, A.,
Jacksoph, S., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpeth, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowls, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshew, L., Louised, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munitas, M., Murphy, M., Nair, L.,
Nankervill, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwakoelemel, O., Okunolu, G., Olarnunsegon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Polndexter, A., Popovic, D., Prims, E., Pu, L.,
Puato, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reib, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, M., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C. D., Smajls, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Vals, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, R., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niedermauser, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.

TITLE
JOURNAL
AUTHORS

REFERENCE

TITLE

AUTHORS

JOURNAL

Sander, W., Saverly, G., Scherer, S., Scott, G., Shateman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Snead, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidl, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usami, K., Valas, R., Vera, V., Villanar, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wodden, H., Worley, K., Wright, D., Wright, R., Wu, U., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 260849)
 Morley, K. C
 Submitted (27-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 260849)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (22-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Sep 22, 2002 this sequence version replaced gi:21746538.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence contigs will be indicated in the feature table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GUPC
 Center clone name: CH230-4702
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 228568 bases at least Q40
 Consensus quality: 232238 bases at least Q30
 Consensus quality: 234535 bases at least Q20
 Estimated insert size: 255181; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_drfc_data.html)
 * NOTE: This sequence may represent more than one clone.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1
 * 20701: contig of 20701 bp in length
 * 20702: 20801: gap of unknown length
 * 20802 41041: contig of 20240 bp in length
 * 41042 41141: gap of unknown length
 * 41142 156928: contig of 115787 bp in length
 * 156929 157028: gap of unknown length
 * 157029 235838: contig of 78810 bp in length
 * 235839 235938: gap of unknown length
 * 235939 260849: contig of 24911 bp in length.
 * Location/Qualifiers
 1..260849
 /organism="Rattus norvegicus"
 FEATURES
 source

```

/misc_feature      /mol type="genomic_DNA"
                   /db xref="taxon:10116"
                   /clone="CH230-4702"
1..1016
   /note="wgs_end_extension
clone_end:T7"
695..1112
   /note="clone_boundary
clone_end:T7
site:EcorI
end_sequence:BH361669"
misc_feature       complement(8162..8721)
                   /note="Clone boundary
clone_end:Sfb
site:EcoRI
end_sequence:BH361671"
misc_feature       96543..98281
                   /note="wgs end_extension
clone_end:Sfp6"

ORIGIN
Alignment Scores:
Pred. No.:          2..8           Length:        260849
Score:              84.00         Matches:        12
Percent Similarity: 100.00%       Conservative:   0
Best Local Similarity: 100.00%    Mismatches:     0
Query Match:        100.00%       Indels:         0
DI:                  2            Gaps:           0

US-09-858-332G-15 (1-12) * AC119447 (1-260849)

CY               1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
Db               CATATCATCATATCATATCATATCATATCATAT 38625
38590 CATATCATATCATATCATATCATATCATATCATATCATAT

RESULT 97
AC102976 LOCUS
DEFINITION Rattus norvegicus clone CH230-3908, *** SEQUENCING IN PROGRESS ***,
AC102976 5 unordered pieces.
ACCESSION AC102976
VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 262405)
Muzny,J.,Marie,, Metzker,M.,Lee,, Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alibekov,S., Amin,A., Angiano,D.,
Anylelechchi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,D., Blankenburg,K., Blyth,P., Brown,M.,
Byrnes,N., Bunay,C., Burch,P., Buttrell,R., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasari,H., Chen,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Cheng,Z., Chu J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Devilla,M.L., Davis,C., Davy-Carroll L., De Anda C., Dedrich,D.,
Dejager,O., Denison,S., Deramo,C., Ding,Y., Dimin,H., Diyya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotte,W., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagge,N., Forbes,L., Foster,M., Foster,P.,
Frieser,C.M., Gabisi,A., Ganter,R., Garcia,A., Garner,T., Garza,M.,
Georgievski,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guervara,W.,
Guarinate,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlik,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hulik,S., Humen,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolyet,A.,
Jolly,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovac,C.,
Karapathy,S., Kraft,C.L., Ledbow,H., Levan,J., Lewis,L., Li Z., Liu J.,
Liu Y., Liu W., Liu Y., London,P., Longacre,S., Lopez,J.,

```

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Lorensbuewa, L., Louisedge, H., Lozado, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartine, M., Mahmood, M., Mallory, K., Mangum, A.,
 Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
 Mathew, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Ming, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwachukwu, O., Okunnu, G., Olarnunpung, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Plankoch, C.,
 Piopfer, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.,
 Puzo, M., Quirio, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojars, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajd, D.,
 Sneed, A., Sodergren, E., Song, X. Z., Stettin, R., Sosa, J.,
 Steindl, M., Strong, R., Sutton, A., Swatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valae, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, D.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczky, R., Woodson, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstein, G., and Gibbs, R. A.

Unpublished
 2 (bases 1 to 262405)
 Direct Submission
 Worley, K. C.
 Direct Submission
 Submitted (24-NOV-2001) Human Genome Sequencing Center, Department

REFERENCE	3 (bases 1 to 262405)
AUTHORS	Rat Genome Sequencing Consortium.
TITLE	Direct Submission
JOURNAL	Submitted (22-SEP-2002) Human Genome Sequencing Center, Department

On Sep 18, 2012 this sequence version replaced g1:22756869. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```

Center:Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GGTm
Center clone name: CH230-39J8
-----
Summary Statistics
Assembly program: Pitrp; version 0.990329
Consensus quality: 192222 bases at least Q40
Consensus quality: 197138 bases at least Q30
Consensus quality: 200012 bases at least Q20
Estimated insert size: 220655; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This is a 'working draft' sequence. It currently
  consists of 5 contigs. The true order of the pieces
  is not known and their order in this sequence record is
  arbitrary. Gaps between the contigs are represented as
  runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
  as soon as it is available and the accession number will
  be preserved.

```

FEATURES	
*	51246: contig of 51246 bp in length
*	1
*	51346: gap of unknown length
*	51347
*	254883: contig of 203537 bp in length
*	254884
*	254984
*	256476: contig of 1493 bp in length
*	256477
*	256576: gap of unknown length
*	256577
*	259335: contig of 2759 bp in length
*	259435: gap of unknown length
*	259436
*	262405: contig of 2970 bp in length.
	Location/Qualifiers

```
misc_feature 14915..80925
              /note="clone_boundary
              clone end:Sp6
              site:BCoRI
              end_sequence:BH285361"
              137042..138655
              /note="wgs_config"
```

Alignment Scores:	
Pred. No.:	2.82
Score:	84.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	2
Length:	262405
Matches:	12
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-858-332G-15 (1-12) X AC102976 (1-262405)

[illegible]

RESULT	98
AC121457	
LOCUS	264615 bp DNA linear HTG 12-OCT-2002
DEFINITION	Rattus norvegicus clone CH230-146J5, *** SEQUENCING IN PROGRESS
ACCESSION	AC121457
VERSION	AC121457.3 GI:23907899
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus

REFERENCE
AUTHORS

1 (bases 1 to 264615)

Mzyny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Aisbrooks, S., Amin, A., Anguiano, D., Ayala-Bechli, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benamed, F., Biewald, K., Blair, J., Blankenbush, K., Blyth, P., Brown, M., Bryant, K., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Derzamo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Dublin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gebisi, A., Gante, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Girdy, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, J., Harvey, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hu, Y. S., Hume, J., Idelbrid, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,


```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 263260: contig of 263260 bp in length
* 263261 263360: gap of unknown length
* 263361 264615: contig of 1255 bp in length.
FEATURES
SOURCE
1.264615
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-146J5"
1..1514
/note="wgs contig"
1565..3140
/note="wgs contig"
4867..6350
/note="wgs contig"
complement(262294..263119)
/note="clone_boundary
clone end:77
site:ECORI
end_sequence:RMBAG51TB"

```

ORIGIN

Alignment Scores:

Area, No.:	2,84	Length:	264615
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-858-332G-15 (1-12) x AC121457 (1-264615)

QY	1	His	Asn	His	Asn	His	Asn	His	Asn	His	Asn	12
Db	161833	C	A	C	A	C	A	C	A	C	A	161868

LOCUS	AC110912.2	268063 bp	DNA	linear	HTG 23-APR-2003
DEFINITION	Mus musculus clone RP24-397B15, WORKING DRAFT SEQUENCE, 68 unmerged pieces.				
ACCESSION	AC110912				
VERSION	AC110912.2	GI:30018053			
KEYWORDS	HTG, HTGS, PHASE1, HTGS, DRAFT.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				

REFERENCE	AUTHORS	TITLE
1	Mammalia, Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus	
2	(Bases 1 to 268063)	
3	Blair, B., Nudbaum, C. and Lander, E.	Mus musculus, clone RP24-397B15
4	Unpublished	
5	2 (Bases 1 to 268063)	
6	Blair, B., Linton, L., Nudbaum, C., Lander, E., Ali, A., Allen, N.,	
7	Anderson, S., Bara, N., Baatien, V., Boguslavsky, L., Boukhalter, B.,	
8	Brown, A., Camrata, J., Campotiano, A., Chang, J., Chazaro, B.,	
9	Choegel, Y., Colangelo, M., Collins, S., Collingmore, A., Cook, A.,	
10	Cooke, P., DeLathau, K., Dewar, K., Diaz, J.S., Dodge, S., Farc, S.,	
11	Ferriz, P., Fitzhugh, W., Gage, D., Galagan, J., Gadda, S.,	
12	Ginde, S., Gold, S., Goyette, M., Graham, L., Grand-pierre, N.,	
13	Hagbo, B., Horton, L., Holme, W., Iliev, I., Johnson, R., Jones, C.,	
14	Kama, A., Karbas, A., Kells, C., Lacroque, K., Lamaizats, R.,	
15	Landers, T., Lebeck, T., Liles, C., Lincogne, K., Linaizats, R.,	

NOTE: Estimated insert size may differ from sequence length.
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

TITLE
JOURNAL
REFERENCE
AUTHORS

MLenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicot, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, U., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talmas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (16-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 268063)

Birren, B., Nusbaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhaiter, B., Camarata, U., Chang, U., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dekrellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hages, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicot, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Rachpaka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 17, 2003 this sequence version replaced gj:18693475.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L2176
Center clone name: 397_B_15

NOTE: This is a 'working draft' sequence. It currently consists of 68 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1
24291 24390: contig of 24290 bp in length
24391 24390: gap of 100 bp
24391 25057: contig of 667 bp in length
25058 25157: gap of 100 bp
25158 25809: contig of 652 bp in length
25810 25909: gap of 100 bp
25910 26753: contig of 844 bp in length
26754 26853: gap of 100 bp
26854 27553: contig of 700 bp in length
27554 29159: contig of 1506 bp in length
29160 29259: gap of 100 bp
29260 30197: contig of 938 bp in length
30198 30297: gap of 100 bp
30298 31460: contig of 1163 bp in length

31461 31560: gap of 100 bp
31561 32590: contig of 1030 bp in length
32591 32690: gap of 100 bp
32691 33703: contig of 1013 bp in length
33704 33803: gap of 100 bp
33804 35074: contig of 1271 bp in length
35075 35174: gap of 100 bp
35175 36345: contig of 1171 bp in length
36346 36445: gap of 100 bp
36446 37894: contig of 1443 bp in length
37895 37994: gap of 100 bp
37995 39318: contig of 1324 bp in length
39319 39418: gap of 100 bp
39419 40661: contig of 1243 bp in length
40662 40761: gap of 100 bp
40762 42129: contig of 1366 bp in length
42130 42229: gap of 100 bp
42230 43797: contig of 1566 bp in length
43798 43897: gap of 100 bp
43899 44862: contig of 965 bp in length
44863 44962: gap of 100 bp
44963 46854: contig of 1892 bp in length
46855 46954: gap of 100 bp
46955 47927: contig of 973 bp in length
47928 48027: gap of 100 bp
48028 49244: contig of 1211 bp in length
49245 49344: gap of 100 bp
49345 50632: contig of 1288 bp in length
50633 50732: gap of 100 bp
50733 52049: contig of 1311 bp in length
52050 52149: gap of 100 bp
52150 53237: contig of 1088 bp in length
53238 53337: gap of 100 bp
53339 55277: contig of 1940 bp in length
55278 55377: gap of 100 bp
55379 56685: contig of 1308 bp in length
56686 56785: gap of 100 bp
56786 58642: contig of 1857 bp in length
58643 58742: gap of 100 bp
58743 60107: contig of 1365 bp in length
60108 60207: gap of 100 bp
60209 61814: contig of 1607 bp in length
61815 61914: gap of 100 bp
61915 63033: contig of 1119 bp in length
63034 63133: gap of 100 bp
63134 64082: contig of 949 bp in length
64083 64182: gap of 100 bp
64183 66519: contig of 2337 bp in length
66520 66619: gap of 100 bp
66620 68694: contig of 2075 bp in length
68695 68794: gap of 100 bp
68795 70253: contig of 1455 bp in length
70254 70353: gap of 100 bp
70354 72323: contig of 1970 bp in length
72324 72423: gap of 100 bp
72424 74360: contig of 1937 bp in length
74361 74460: gap of 100 bp
74461 76252: contig of 1792 bp in length
76253 76352: gap of 100 bp
76353 78129: contig of 1777 bp in length
78130 78229: gap of 100 bp
78230 80637: contig of 2408 bp in length
80638 80737: gap of 100 bp
80738 82870: contig of 2133 bp in length
82871 82970: gap of 100 bp
82971 85712: contig of 2742 bp in length
85713 85812: gap of 100 bp
85813 87890: contig of 2078 bp in length
87891 87990: gap of 100 bp
87991 90163: contig of 2173 bp in length
90164 90263: gap of 100 bp
90264 92559: contig of 2296 bp in length
92560 92659: gap of 100 bp

```

* 92660 95748: contig of 3089 bp in length
* 95749 95848: gap of 100 bp
* 95849 98374: contig of 2526 bp in length
* 98375 98474: gap of 100 bp
* 98475 101777: contig of 3303 bp in length
* 101778 101877: gap of 100 bp
* 101878 104813: contig of 2936 bp in length
* 104814 104913: gap of 100 bp
* 104914 107515: contig of 2602 bp in length
* 107516 110328: gap of 100 bp
* 110329 110428: contig of 2713 bp in length
* 110429 114176: gap of 100 bp
* 114177 114276: contig of 3748 bp in length
* 114277 118279: gap of 100 bp
* 118280 118379: contig of 4003 bp in length
* 118380 122029: gap of 100 bp
* 122030 122129: contig of 3650 bp in length
* 122130 126278: gap of 100 bp
* 126279 126378: contig of 4149 bp in length
* 126379 130335: gap of 100 bp
* 130336 130435: contig of 3957 bp in length
* 130436 134925: gap of 100 bp
* 134926 135025: contig of 4490 bp in length
* 135026 138675: gap of 100 bp
* 138676 138775: contig of 3650 bp in length
* 138776 143086: gap of 100 bp
* 143087 143086: contig of 4311 bp in length

```

Alignment Scores:

```

Pred. No.: 2.88 Length: 268063
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

```

US-09-858-332G-15 (1-12) x AC110912 (1-268063)

Qy 1 HisAenHisAenHisAenHisAenHisAenHisAen 12

Db 250744 CATATATCATATATCATATATCATATATCATATAT 250709

RESULT 100

AC091323

LOCUS

DEFINITION

Mus musculus chromosome 17 clone RP23-333B5 map 17, *** SEQUENCING

ACCESSION

AC091323.4 GI:39752830

KEYWORDS

HTG; HTGS PHASE1; HTGS FULLTOP; HTGS_ACTIVEFIN.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

```

Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S.,
Severy,P., Sougnier,C., Spencer,B., Strange-Thomann,N., Seaman,S.,
Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Testaye,S.,
Theodore,J., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,M.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (15-APR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 268644)
JOURNAL
REFERENCE
AUTHORS
Birren,B., Nuebaum,C., Lander,E., Abouelella,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Boukhaltier,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K., Dewar,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Halez,N., Hagopian,D., Hages,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lander,E., Levine,R., Maclean,C.,
Macdonald,P., Major,J., Manning,V., Matthews,C., McCarthy,M.,
Meldrum,J., Meneus,L., Mihova,T., Muller,R., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunhhang,P., Pierre,N.,
Rachupka,A., Ramsamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (03-MAR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 12, 2003 this sequence version replaced gi:24182393.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L13287
Center clone name: 333_E_5

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 82781: contig of 82781 bp in length
* 82782 82881: gap of 100 bp
* 82882 126562: contig of 4381 bp in length
* 126563 126662: gap of 100 bp
* 126663 134445: contig of 7783 bp in length
* 134446 134545: gap of 100 bp
* 134546 142101: contig of 7556 bp in length
* 142102 142201: gap of 100 bp
* 142202 203346: contig of 6145 bp in length
* 203347 203446: gap of 100 bp
* 203447 235687: contig of 32241 bp in length
* 235688 235787: gap of 100 bp
* 235788 251563: contig of 15776 bp in length
* 251564 251663: gap of 100 bp
* 251664 268644: contig of 16981 bp in length.
Location/Qualifiers
1. 268644
/organism="Mus musculus"

```

FEATURES

source

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: October 4, 2005, 11:23:56 ; Search time 426 Seconds
(without alignments)
166.753 Million cell updates/sec

Title: US-09-858-332G-15
Perfect score: 84
Sequence: 1 NHHNHHNHHNHN 12

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=x1p
-Q=/cgrn2.1/USFTO.spool.p/US09858332/runat_04102005_122256_4283/app_query.fasta_1.199
-DB=N Geneseq_16Dec04 -QFMT=faalap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNIT9=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=100 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=100
-MODE=LOCAL -OUTFMT=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09858332 @CGN 1.1 708 @runat_04102005_122256_4283 -NCPU=6 -ICPU=3
-NO MAP -LARGEOUTPUT -NEG_SCORES=0 -WAIT -DSPELLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WAPN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FPAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: Geneseq_16Dec04:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84	100.0	36	8	ABX94280 DNA varia
2	84	100.0	42	8	ABV76005 Hexa-hist
3	84	100.0	100	8	ABV76004 Splice do
4	84	100.0	115	10	ADD08922 Bovine ge
5	84	100.0	519	10	ADB58733 Toxicity-

6	84	100.0	519	10	ADB53430	ADB53430 Primary r
7	84	100.0	1897	2	AAQ28734	AAQ28734 Sequence
8	84	100.0	1897	2	AAE4394	AAE4394 Saccharom
9	84	100.0	1898	6	ABQ76384	ABQ76384 S. cerevi
10	84	100.0	4938	8	ABV75997	ABV75997 Donor vec
11	84	100.0	6525	8	ABV75999	ABV75999 Donor vec
12	84	100.0	7487	8	ABV76000	ABV76000 Vector pl
13	84	92.9	139	10	ADD08924	ADD08924 Plasmid p
14	78	92.9	205	10	ADD08921	ADD08921 Bovine ge
15	78	92.9	587	10	ADD08920	ADD08920 Bovine VD
16	78	92.9	1680	10	ACF68722	ACF68722 Phototrab
17	78	92.9	1731	12	ADL02680	ADL02680 DNA encod
18	78	92.9	110000	10	ACF67367_15	ACF67367_15 Phototrab
19	78	92.9	249878	10	ACF65381	ACF65381 Phototrab
20	77	91.7	537	13	ACN56255	ACN56255 Cotton an
21	77	91.7	16489	4	AAE27823	AAE27823 DNA encod
22	77	91.7	16489	10	ADB94626	ADB94626 Novel hum
23	76	90.5	1154	5	AAAD1057	AAAD1057 Lama HCV
24	76	90.5	1172	5	AAAD1058	AAAD1058 Lama HCV
25	76	90.5	1521	4	AAE23387	AAE23387 Candida a
26	76	90.5	1521	6	ABZ31758	ABZ31758 Candida a
27	76	90.5	4555	4	AAE55447	AAE55447 Nucleotid
28	76	90.5	4565	8	ABX35442	ABX35442 Transpos
29	76	90.5	4810	8	ABX35441	ABX35441 DNA encod
30	73	86.9	862	3	AAE36337	AAE36337 Arabidops
31	73	86.9	862	3	AAE36056	AAE36056 Arabidops
32	73	86.9	960	3	AAE52042	AAE52042 Arabidops
33	72	85.7	3042	6	ABQ76560	ABQ76560 C. albica
34	72	85.7	5826	6	ABL32563	ABL32563 Human imm
35	72	85.7	6236	6	ABL33630	ABL33630 Human imm
36	72	85.7	6236	6	ABN80202	ABN80202 Human che
37	71	84.5	74037	6	ABK94412	ABK94412 DNA encod
38	70	83.3	904	3	AAE26665	AAE26665 Candida a
39	70	83.3	2556	2	AAV19873	AAV19873 Drosophi
40	70	83.3	5511	3	AAAE1847	AAAE1847 Cryptospo
41	70	83.3	5511	6	ABT04776	ABT04776 C parvum
42	70	83.3	7334	3	AAAE1846	AAAE1846 Cryptospo
43	70	83.3	7334	6	ABT04775	ABT04775 C parvum
44	70	83.3	24076	6	ABO76608	ABO76608 C. albica
45	69	82.1	3855	4	ABL30105	ABL30105 Drosophi
46	69	82.1	10907	4	ABL30104	ABL30104 Drosophi
47	68.5	81.5	2094	4	ABL15969	ABL15969 Drosophi
48	68.5	81.5	4569	4	ABL15968	ABL15968 Drosophi
49	67	79.8	315	6	ABN67727	ABN67727 Streptoco
50	67	79.8	4779	4	ABL17932	ABL17932 Drosophi
51	67	79.8	110000	6	ABN71527_20	ABN71527_20 Continuation (21 o
52	66	78.6	1338	12	ADL42533	ADL42533 Plant tra
53	66	78.6	1338	12	ADL42533	ADL42533 Soybean o
54	66	78.6	5826	6	ABL32562	ABL32562 Human imm
55	66	78.6	6236	6	ABL33631	ABL33631 Human imm
56	66	78.6	6236	6	ABN80203	ABN80203 Human che
57	66	78.6	28360	4	ABL06142	ABL06142 Drosophi
58	65	77.4	334	6	ABE69323	ABE69323 Novel mut
59	65	77.4	599	10	ADD35154	ADD35154 Mouse mit
60	65	77.4	700	4	AAE23528	AAE23528 Mouse NAD
61	65	77.4	927	4	ABL27329	ABL27329 Drosophi
62	65	77.4	1958	4	ABL27328	ABL27328 Drosophi
63	65	77.4	5163	2	AAV20700	AAV20700 Cryptospo
64	65	77.4	5163	3	AAAE1849	AAAE1849 ORF encod
65	65	77.4	5163	6	ABT04778	ABT04778 C parvum
66	65	77.4	5318	2	AAV20701	AAV20701 Cryptospo
67	65	77.4	5318	3	AAAE1848	AAAE1848 DNA encod
68	65	77.4	5318	6	ABT04777	ABT04777 C parvum
69	65	77.4	13033	4	ABL14054	ABL14054 Drosophi
70	65	77.4	110000	6	ABO67196_5	ABO67196_5 Continuation (28 o
71	65	77.4	110000	6	ABO69245_27	ABO69245_27 Continuation (27 o
72	65	77.4	110000	6	ABA03041_26	ABA03041_26 Continuation (27 o
73	64	76.2	65	6	ABE227352	ABE227352 Candida e
74	64	76.2	148	4	AAI21768	AAI21768 Probe #11
75	64	76.2	148	4	ABAE66942	ABAE66942 Human foe
76	64	76.2	148	4	AAI47050	AAI47050 Probe #15
77	64	76.2	148	4	ABAE48927	ABAE48927 Human bre
78	64	76.2	148	4	ABA33909	ABA33909 Probe #12

79	64	76.2	148	4	AAK40999
80	64	76.2	148	4	AAK15273
81	64	76.2	148	4	ABSA0587
82	64	76.2	148	5	AA107454
83	64	76.2	148	6	ABSI14969
84	64	76.2	254	10	ADB75735
85	64	76.2	381	4	AA112579
86	64	76.2	381	4	ABSA4279
87	64	76.2	381	4	AA133926
88	64	76.2	381	4	ABSA3326
89	64	76.2	381	4	ABSA24034
90	64	76.2	381	4	AAK27997
91	64	76.2	381	4	AAK02558
92	64	76.2	381	5	ABSA27588
93	64	76.2	381	5	ABSA02467
94	64	76.2	381	6	ABSA02467
95	64	76.2	444	6	ABSA02467
96	64	76.2	479	4	AA110862
97	64	76.2	479	4	ABSA2513
98	64	76.2	479	4	AA132122
99	64	76.2	479	4	ABSA42091
100	64	76.2	479	4	ABSA22302

ALIGNMENTS

RESULT 1
ABX94280
ID ABX94280 standard; DNA; 36 BP.
XX
XX
AC ABX94280;
XX
DT 17-JUN-2003 (first entry)
XX
DE DNA variant sequence encoding affinity purification peptide #2.
XX

XX Metal ion affinity peptide: fusion protein; protein purification;
KM metal chelating resin; metal ion chelate resin; hard metal; Fe3+; Ca2+;
KM Al3+; Co2+; Cu2+; Ni2+; Zn2+; Co2+; protein-protein interaction;
KM metal ion affinity peptide-tagged recombinant protein; IMAC;
KM DNA-protein interaction; immobilised metal ion affinity chromatography;
KM gene expression; phosphorylation state; ds.
XX

OS Synthetic.

XX Key Location/Qualifiers
FH 1..36
FT /*tag= a
FT /partial
FT /note= "This sequence lacks both start and stop codons"

XX US2002164718-A1.
XX
XX 07-NOV-2002.
XX
XX 15-MAY-2001; 2001US-00858332.
XX

XX 25-SEP-1998; 98US-0101867P.
XX 23-SEP-1999; 99US-00404017.
XX

XX (TCHAG/) TCHAGA G S.
XX (JOKH/) JOKHADZE G G.
XX

XX Tchaga GS, Jokhadze GG;
XX

XX WPI: 2003-361747/34.
XX P-PSDB; ABU08447.
XX

XX New metal ion affinity peptide useful, when fused to a fusion partner
PT polypeptide, for protein purification methods and to study protein-
PT protein interactions and nucleic acid-protein interactions.
XX

PS Example 1; Fig 3; 23pp; English.

XX The present invention relates to metal ion affinity peptides, fusion
CC proteins containing metal ion affinity peptides, and polynucleotide
CC sequences encoding the fusion proteins. The presence of a metal ion
CC affinity peptide in a fusion protein allows purification of the fusion
CC protein on a metal chelating resin. The method involves contacting a
CC sample comprising a fusion protein with a metal ion chelate resin
CC comprising a first metal ion, preferably a hard metal ion such as Fe3+,
CC Ca2+ or Al3+ and eluting any resultant bound fusion protein from the
CC resin. The resin comprises an immobilised Co2+ ion. The method further
CC comprises contacting the sample with a second immobilised metal ion
CC affinity resin comprising a second immobilised metal ion and eluting any
CC resultant bound fusion protein from the first and second resins. The
CC second metal ion is an intermediate metal ion such as Cu2+, Ni2+, Zn2+ or
CC Co2+. The metal ion affinity peptide-tagged recombinant proteins are
CC useful for the study of protein-protein interactions and nucleic acid
CC molecule-protein interactions, using solid phase immobilised metal ion
CC affinity chromatography (IMAC). They are also useful in high throughput
CC systems which find use in massive parallel gene expression experiments
CC e.g. to determine the effect of an agent on synthesis of a protein or set
CC of proteins, to analyse developmental stage-specific, or tissue-specific
CC synthesis of a protein and to analyse the phosphorylation state of a
CC protein. These methods find use in applications to characterise a protein
CC of unknown identity or function, and in enzymatic reactions. ABX94279-
CC ABX94283 represent DNA variants that encode affinity purification
CC peptides

XX SQ Sequence 36 BP; 18 A, 12 C, 0 G, 6 T, 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Matches:
Score: 84.00	36	12
Percent Similarity: 100.00%	Conservative: 0	
Best Local Similarity: 100.00%	Mismatches: 0	
Query Match: 100.00%	Indels: 0	
DB: 8	Gaps: 0	

US-09-858-332G-15 (1-12) x ABX94280 (1-36)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
Db 1 CATTAACATTAACATTAACATTAACATTAAC 36

RESULT 2
ABV76005
ID ABV76005 standard; DNA; 42 BP.
XX
XX
AC ABV76005;
XX
DT 11-FEB-2003 (first entry)
XX

XX Hexa-histidine-asparagine tag peptide coding sequence.
XX
XX Vector; splice donor; recombinase; intron; affinity tag; gene; ds.
XX

OS Synthetic.

XX Key Location/Qualifiers
FH 1..42
FT /*tag= a
FT /product= "6xHis tag"
FT /note= "the CDS does not include a start or stop codon"

XX WO200283910-A2.
XX

XX 24-OCT-2002.
XX

XX 17-JAN-2002; 2002MO-US001604.
XX

XX 18-JAN-2001; 2001US-0263358P.
XX

XX (CLON-) CLONTECH LAB INC.
XX

[illegible]

```

PR 18-JAN-2001; 2001US-0263358P.
XX
XX (CLON-) CLONTECH LAB INC.
XX
XX Farmer AA;
XX
DR WP1; 2003-075553/07.
XX
PS Example 2; Page 42; 64pp; English.
XX
CC The present sequence is that of a splice donor-containing oligonucleotide
CC that was cloned into vector pDNR-1 (see ABV75997) at ApaI and AvrII
CC sites. The oligonucleotide was designed to place stop codons (TAG and
CC TGA) in the 2 unused reading frames present in the multiple cloning site
CC of pDNR-1. A 6xHis tag sequence was also included to facilitate protein
CC purification in bacteria. The construct can be used in the method of the
CC invention. The invention provides methods for producing a vector that
CC includes at least one splicing intron. The vector is produced from
CC donor and acceptor vectors that each include a site-specific recombinase
CC site, as well as splice donor and acceptor sites that, upon site-specific
CC recombination of the donor and acceptor vectors, define an intron in the
CC product vector of the recombination. Examples of donor, acceptor and
CC product vectors are given in ABV75997-6000. The method is useful for
CC production of vectors encoding C-terminal tagged fusion proteins, and
CC expression vectors encoding pure proteins. Rapid transfer of a DNA
CC molecule from one vector to another is achieved in vitro or in vivo
CC without the need to rely upon restriction enzyme digestions
XX
SQ Sequence 100 BP; 33 A; 17 G; 24 C; 26 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 0.00211 Length: 100
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
XX
US-09-858-332G-15 (1-12) x ABV76004 (1-1100)
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
Db |||||
61 CATATCATATCATATCATATCATATCATATCATATCATATCAAC 96
RESULT 4
ADD08922/c ADD08922 standard; DNA; 115 BP.
XX
AC ADD08922;
XX
DT 01-JAN-2004 (first entry)
DE Bovine germ-line D nucleic acid sequence D-2 SEQ ID NO:3.
XX
XX bovine; VDJ gene cassette; bovine germ-line; vaccine; immunotherapy;
KW immunoglobulin; gene; ds.
XX
OS Synthetic.
OS Bos taurus.
XX
PN CA2382766-A1.
XX
PD 20-OCT-2002.
XX
PF 19-APR-2002; 2002CA-02382766.
XX
PR 20-APR-2001; 2001US-0284899P.
XX

```

PA (SAINI/) SAINI S. S.
 PA (KAUS/) KAUSHIK A.
 XX
 PI Saini SS, Kaushik A;
 XX
 DR WPI, 2003-31427/31.
 XX
 PT Novel isolated bovine VDJ gene cassette useful for preparing a vaccine
 XX vector, and in immunotherapy.
 XX
 PS Claim 5; Fig 2; 29pp; English.
 XX
 CC The present invention describes an isolated bovine VDJ gene cassette (I)
 CC comprising the 587 nucleotide sequence of ABD08920. Also described: (1) a
 CC recombinant plasmid pBFIH-24 (II); (3) an isolated bovine germine D
 CC sequence (III) comprising the sequence of D-1 (ABD08921), D-2 (ABD08922)
 CC and/or D-3 (ABD08923); (4) a recombinant plasmid pBD148; and (5) a
 CC recombinant vector, in immunotherapy. (I) or (II) can be used for preparing a
 CC vaccine vector, (I) provides the novel ability to develop chimeric
 CC immunoglobulin molecule capable of incorporating both linear T cell
 CC epitope(s) (CDR1H and CDR2H) as well as conformational B cell epitope(s)
 CC (exceptionally long CDR3H). As a result, the germine encodes BFIH VDJ
 CC cassette isolated from bovine foetal B cells is suitable for
 CC immunoglobulin antigenisation with both B and T cell epitopes and use as
 CC vaccine vector. Further multiple epitopes can be incorporated for
 CC development of multivalent vaccine by replacing at least a portion of an
 CC immunoglobulin molecule with the desired epitope such that the functional
 CC ability of both epitope(s) and parent VDJ rearrangement is retained. The
 CC antigenised immunoglobulin incorporating both T and B cell epitopes of
 CC interest is useful for development of oral vaccines for humans apart from
 CC other species including cattle. (III) provides additional opportunities
 CC for sustaining the capacity for antibody diversification in cattle
 CC essential for immunocompetence by selective breeding strategies that
 CC incorporate immunoglobulin gene markers. (III) is unique to cattle and
 CC therefore are useful in forensic analysis. The VDJ cassette isolated from
 CC cattle provides a natural source of immunoglobulin that is capable of
 CC incorporating multiple, both B and T cell epitopes, and is, therefore,
 CC most suitable as vaccine vector across species by using species-specific
 CC different isotypes, as or if required.
 XX
 SQ Sequence 115 BP; 23 A; 16 C; 35 G; 41 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 0.00243 Length: 115
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0
 US-09-858-332g-15 (1-12) x ABD08922 (1-115)
 QY 1 H1sAsnH1sAsnH1sAsnH1sAsnH1sAsnH1sAsn 12
 Db 80 CATTAACCATTAACCATTAACCATTAACCATTAAC 45
 RESULT 5
 ADBS8733
 ID ADBS8733 standard; DNA; 519 BP.
 XX
 AC ADBS8733;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Toxicity-related gene, SEQ ID 3759.
 XX
 KW Toxic; toxin; gene expression profile; hepatotoxicity; liver;
 KW drug screening; toxicity assay; ds.
 XX
 OS Unidentified.
 XX
 FN WO2003064624-A2.

XX
 PD 07-AUG-2003.
 XX
 PF 31-JAN-2003; 2003WO-US003194.
 XX
 PR 31-JAN-2002; 2002US-0060087.
 XX
 PR 15-MAR-2002; 2002US-0364045P.
 XX
 PR 15-MAR-2002; 2002US-0364055P.
 XX
 PR 30-DEC-2002; 2002US-0436643P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M,
 XX WPI, 2003-689530/65.
 DR
 XX
 PT Predicting a toxic effect of a compound, useful in identifying toxicity
 PT markers in liver tissues or cells for drug screening and toxicity assays,
 PT comprises preparing gene expression profile of tissue or cells exposed to
 XX the compound.
 XX
 PS Claim 1; SEQ ID NO 3759; 1156pp; English.
 XX
 CC The present invention relates to a method for predicting a toxic effect
 CC of a compound. The method comprises preparing a gene expression profile
 CC of a tissue or cell sample exposed to the compound, and comparing the
 CC gene expression profile to a database comprising SEQ ID 1-4925, where
 CC differential expression of the gene indicates at least one toxic effect.
 CC The method is useful for predicting at least one toxic effect of a
 CC compound, predicting hepatotoxicity or the progression of a toxic effect
 CC of a compound, identifying an agent that modulates the onset or
 CC progression of a toxic response, predicting the cellular pathways that a
 CC compound modulates in a cell, and identifying an agent that modulates at
 CC least one activity of a protein. The method and compositions of the
 CC present invention using a database of genes having liver toxin-induced
 CC differential expression, are useful in identifying toxicity markers in
 CC liver tissues or cells for drug screening and toxicity assays. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 519 BP; 142 A; 130 C; 136 G; 111 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 0.0111 Length: 519
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0
 US-09-858-332g-15 (1-12) x ADBS8733 (1-519)
 QY 1 H1sAsnH1sAsnH1sAsnH1sAsnH1sAsnH1sAsn 12
 Db 128 CACACCATTAACCATTAACCATTAACCATTAAC 163
 RESULT 6
 ADBS3430
 ID ADBS3430 standard; DNA; 519 BP.
 XX
 AC ADBS3430;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3972.
 XX
 KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
 KW toxicity marker; toxicity progression; drug screening;
 KW primary rat hepatocyte toxicity modelling; gene; ds.
 XX
 OS Rattus norvegicus.
 XX

CC of three OP11 mutant alleles - OP11-1 (JHO-6D), OP11-2 (NO80) and OP11-3
CC (NO99) identifies nonsense mutations in each of the three alleles, all
CC located within the first polypeptide stretch of Aa. Alleles OP11-3 and
CC OP11-2 have a TAA codon located at nucleotide 1294 and 1312 respectively,
CC whereas OP11-1 has a TAG codon at nucleotide 1315. (Updated on 25-MAR-
CC 2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS field.)
XX

Sequence 1897 BP; 592 A; 403 C; 464 G; 438 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.041	Length:	1897
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-858-332G-15 (1-12) x AAQ28734 (1-1897)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
DB 1730 CATATCATATCATATCATATCATATCATATCATAT 1765

RESULT 8

ID AAT64394 standard; DNA; 1897 BP.

AC AAT64394;

DT 25-MAR-2003 (revised)

DT 21-MAY-1997 (first entry)

DE Saccharomyces cerevisiae OP11 gene.

KM Negative regulator of phospholipid biosynthesis; Op11 protein;
KM leucine zipper; polyglutamine tract; inositol-1-phosphate; ds.
XX

OS Saccharomyces cerevisiae.

XX Key Location/Qualifiers

FT CDS 439..1653

FT /*tag= a

FT /transl_except= pos:1291..1293, aa:Gln

FT /product= "Op11p"

XX US5599701-A.

PD 04-FEB-1997.

PF 30-MAY-1995; 95US-00453461.

PR 14-JAN-1991; 91US-00640495.

PR 02-DEC-1993; 93US-00165182.

XX (UYCA-) UNIV CARNEGIE MELLON.

XX White MJ, Henry SA;

XX WPI, 1997-118296/11.

XX P-PSDB; AAW14909.

PT Recombinant yeast with enhanced inositol prodn. - lacking OP11 gene.

XX Claim 1; Fig 8; 28pp; English.

CC A genetically engineered Saccharomyces yeast cell in which all copies of
CC the OP11 gene of the present sequence have been deleted is claimed. The
CC genetically engineered yeast cell is useful for enhanced production of
CC inositol, inositol-containing metabolites or phospholipids, including myo
CC -inositol and inositol-1-phosphate, for human or animal consumption.
CC (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 1897 BP; 593 A; 404 C; 462 G; 438 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.041	Length:	1897
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-858-332G-15 (1-12) x AAT64394 (1-1897)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
DB 1730 CATATCATATCATATCATATCATATCATATCATAT 1765

RESULT 9

ID ABQ76384 standard; cDNA; 1898 BP.

AC ABQ76384;

DT 21-NOV-2002 (first entry)

DE S. cerevisiae BAX-associated cDNA fragment SEQ ID 193.

KM Bax; Bax-resistance; cytosolic; fungicide; immunosuppressive; virucide;
KM vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
KM apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
KM neurodegeneration; cell death; ss.

XX Saccharomyces cerevisiae.

XX WO200264766-A2.

XX 22-AUG-2002.

XX 21-DEC-2001; 2001WO-EP015398.

XX 22-DEC-2000; 2000EP-00870318.

XX 04-JAN-2001; 2001EP-00870002.

XX 09-JAN-2001; 2001EP-00870003.

XX (JNANC) JANSSEN PHARM NV.

XX Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;

XX WPI, 2002-667002/71.

XX P-PSDB; ABG93118.

PT New isolated nucleic acid representing a synthetic BAX-gene, useful as
PT medicament for treating, preventing and/or alleviating yeast or fungal
PT infections or proliferative disorders, or for preventing apoptosis in
PT certain diseases.

XX Claim 36; Fig 1; 344pp; English.

CC This invention describes a novel nucleic acid representing a synthetic
CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
CC resistant yeast or fungi, identifying, or obtaining and identifying
CC Candida spp. sequences that are differentially expressed in a pathway
CC eventually leading to programmed cell death or identifying inhibitors or
CC inhibitor sequences of Bax-induced cell death. The products of the
CC invention have cytostatic, fungicide, immunosuppressive, virucide and
CC vasotropic activity and can be used in vaccines or for gene therapy. The
CC isolated nucleic acids, polypeptides, pharmaceutical compositions,
CC antisense molecules and antibodies are useful as medicaments or in
CC preparing a medicament for treating, preventing and/or alleviating
CC diseases associated with yeast or fungi or proliferative disorders, such
CC as cancer, or for preventing apoptosis in certain diseases. The compounds
CC or polypeptides, or the genetically modified organism are useful for
CC preparing a medicament for modifying the endogenous flora of humans and
CC other mammals. The vaccine is useful for immunising against yeast or
CC fungal infections. Apoptosis-related diseases include autoimmune disease,

CC ischaemia, diseases related with viral infections or neurodegenerations.
CC This sequence represents a polynucleotide associated with the Bax gene
CC described in the disclosure of the invention

XX Sequence 1898 BP; 613 A; 406 C; 360 G; 519 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.041	Length:	1898
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-858-332G-15 (1-12) x ABQ76384 (1-1898)

Qy 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 195 CATATCATATCATATCATATCATATCATATCATAT 230

RESULT 10

ID ABV75997 standard; DNA; 4938 BP.

XX ABV75997;

DT 11-FEB-2003 (first entry)

XX Donor vector pDNR-Dual.

XX Vector; pDNA-Dual; recombinase; intron; ds.

XX Unidentified.

PN WO200283910-A2.

PD 24-OCT-2002.

PF 17-JAN-2002; 2002WO-US001604.

PR 18-JAN-2001; 2001US-0263358P.

PA (CLON-) CLONTECH LAB INC.

PI Farmer AA;

DR WPI; 2003-075553/07.

XX New sequence specific recombinase based system comprising a donor and an
XX acceptor vector having at least one splice site, where each vector
XX comprises at least one recombination site, useful in preparing an intron-
XX containing vector.

XX Example 1; Page 34-36; 64pp; English.

XX The present sequence is that of donor vector pDNR-dual, which includes a
XX chloramphenicol-resistance open reading frame (ORF) flanked by loxP
XX sites. The donor vector was used in an example of the invention in order
XX to test the utility of intron-splicing to enable tagging of a protein of
XX interest in a donor vector with a peptide tag or protein in an acceptor
XX vector. A luciferase gene was cloned into the multiple cloning site of
XX pDNA-dual, creating vector pDNR-Dual-Luc (see ABV75999). pDNR-Dual-Luc
XX and acceptor vector pLPS-BGFP (see ABV75998) were recombined in vitro
XX using Cre recombinase, creating vector pLPS-Luc-BGFP (see ABV76000). This
XX vector has a splice donor sequence from the donor vector and a splice
XX acceptor sequence from the acceptor vector, together creating an
XX artificial intron between the 3' end of the luciferase gene and the 5'
XX end of an EGFP tag sequence. This intron was composed of the
XX chloramphenicol-resistance ORF, a loxP site and an ampicillin promoter
XX sequence. The construct generated a properly spliced mRNA in transfected
XX HEK293 cells, so enabling expression of a luciferase EGFP fusion protein.
XX This is an example of the method of the invention for producing intron-
XX containing vectors using donor and acceptor vectors and a sequence-

CC specific recombinase. The method is useful for production of vectors
CC encoding C-terminal tagged fusion protein, and expression vectors
CC encoding pure protein. Rapid transfer of a DNA molecule from one vector
CC to another is achieved in vitro or in vivo without the need to rely upon
CC restriction enzyme digestions

XX Sequence 4938 BP; 1492 A; 1122 C; 1091 G; 1233 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.108	Length:	4938
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	8	Gaps:	0

US-09-858-332G-15 (1-12) x ABV75997 (1-4938)

Qy 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 141 CATATCATATCATATCATATCATATCATATCATAC 176

RESULT 11

ID ABV75999 standard; DNA; 6525 BP.

XX ABV75999;

DT 11-FEB-2003 (first entry)

XX Donor vector pDNR-Dual-Luc.

XX Vector; pDNA-Dual-Luc; luciferase; enzyme; firefly; recombinase; intron;
XX ds.

XX Photinus.

OS Unidentified.

OS Chimeric.

PN WO200283910-A2.

PD 24-OCT-2002.

PF 17-JAN-2002; 2002WO-US001604.

PR 18-JAN-2001; 2001US-0263358P.

PA (CLON-) CLONTECH LAB INC.

PI Farmer AA;

DR WPI; 2003-075553/07.

XX New sequence specific recombinase based system comprising a donor and an
XX acceptor vector having at least one splice site, where each vector
XX comprises at least one recombination site, useful in preparing an intron-
XX containing vector.

XX Example 1; Page 33-39; 64pp; English.

XX The present sequence is that of donor vector pDNR-dual-Luc, which
XX includes a chloramphenicol-resistance open reading frame (ORF) and a
XX firefly luciferase gene flanked by loxP sites. The donor vector was used
XX in an example of the invention in order to test the utility of intron-
XX splicing to enable tagging of a protein of interest in a donor vector
XX with a peptide tag or protein in an acceptor vector. pDNR-Dual-Luc and
XX acceptor vector pLPS-BGFP (see ABV75998) encoding an EGFP tag were
XX recombined in vitro using Cre recombinase, creating vector pLPS-Luc-BGFP
XX (see ABV76000). This vector has a splice donor sequence from the donor
XX vector and a splice acceptor sequence from the acceptor vector, together
XX creating an artificial intron between the 3' end of the luciferase gene
XX and the 5' end of an EGFP tag sequence. This intron was composed of a
XX chloramphenicol-resistance open reading frame, a loxP site and an

Pred. No.:	0.143	Length:	6525
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	A	Genes:	0

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 1728 CATATCATATCATATCATATCATATCATACAC 1763

ABV76000
ID ABV76000 standard; DNA; 7487 BP.
XX

Accession	Gene	Accession	Gene
AA	Vector	DE	Vector
DE	plpS-Luc-EGFP.	DE	plpS-Luc-EGFP.
XX	Vector	XX	Vector
KN	plpS-Luc-EGFP	KN	plpS-Luc-EGFP
KN	Luciferase; enzyme: firefly; recombinase; intron	KN	Luciferase; enzyme: firefly; recombinase; intron
KN	ds.	KN	ds.

WO200283910-A2.

17-JAN-2002; 2002WO-US001604.

(CLON-) CLONTECH LAB INC.

WPI; 2003-075553/07.

New sequence specific recombinase based system comprising a donor and an acceptor vector having at least one splice site, where each vector comprises at least one recombination site, useful in preparing an intron containing vector.

Example 1; Page 39-41; 64pp; English.

The present sequence is that of vector pRBS-luc-Egfp, which includes an artificial intron including a chloramphenicol resistance (Cmr) open reading frame, a loxp site, and an ampicillin resistance gene promoter. The intron is placed between the 3' end of a luciferase gene and the 5' end of an EGFP tag sequence. It was created by recombination *in vitro* using Cre recombinase and donor vector pDNR-Dual-luc (see ABV75999), which included the luciferase gene, Cmr open reading frame and ampicillin promoter and acceptor vector pRBS-EGFP (see ABV75998), which included the EDFP sequence and a single loxp site. The luciferase gene was placed such

XX Sequence 7487 BP; 1951 A; 1871 C; 1860 G; 1805 T; 0 U; 0 Other

Alignment Scores:	
Pred. No.:	
Score:	0.164
Percent Similarity:	84.00
Best Local Similarity:	100.00%
Query Match:	100.00%
8:	
8	
Gaps:	0
Length:	7487
Matches:	12
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-858-332G-15 (1-12) x ABV76000 (1-7487)

1 HisaBnHisaBnHisaBnHisaBnHisaBn 12

Db 2316 CATAATCATAATCATAATCATAATCACAAC 2351

RESULT 13
ADD08924/c
ID ADD08924 standard; DNA; 139 BP.
Y

ADD08924;
01-JAN-2004 (first entry)

Plasmid pBF1H1-D3 nucleic acid sequence SEQ ID NO:5.

bovine; VDJ gene cassette; bovine germline; vaccine; immunotherapy; immunoglobulin; gene; ds.

Synthetic.
Bos taurus.

CA2382766-A1

20-OCJ-2002.

19-APR-2002; 2002CA-02382766.

20-APR-2001; 2001US-0284899P.

(SAIN/) SAINI S S.
(KAUS/) KAUSHIK A.

Saini SS, Kaushik A;

WPI; 2003-314127/31.

Novel isolated bovine VDJ gene cassette useful for preparing a vaccine vector, and in immunotherapy.

Claim 7; Fig 3; 29pp; English.

The present invention describes an isolated bovine VDJ gene cassette (I) comprising the 587 nucleotide sequence of AbDO8920. Also described: (1) a recombinant plasmid pBfH1-24 (II); (3) an isolated bovine germline D sequence (III) comprising the sequence of D-1 (AbDO6821), D-2 (AbDO8922) and/or D-3 (AbDO8923); (4) a recombinant plasmid pBfG145; and (5) a recombinant plasmid pBfH1-D3. (I) or (II) can be used for preparing a

CC vaccine vector, in immunotherapy. (III) is useful as a molecular marker
CC and as a DNA probe. (I) provides the novel ability to develop chimeric
CC immunoglobulin molecule capable of incorporating both linear T cell
CC epitope(s) (CDR1H and CDR2H) as well as conformational B cell epitope(s)
CC (exceptionally long CDR3H). As a result, the germline encodes Bf1H1 VDJ
CC cassette isolated from bovine foetal B cells is suitable for
CC immunoglobulin antigenisation with both B and T cell epitopes and use as
CC vaccine vector. Further multiple epitopes can be incorporated for
CC development of multivalent vaccine by replacing at least a portion of an
CC immunoglobulin molecule with the desired epitope such that the functional
CC ability of both epitope(s) and parent VDJ rearrangement is retained. The
CC antigenised immunoglobulin incorporating both T and B cell epitopes of
CC interest is useful for development of oral vaccines for humans apart from
CC other species including cattle. (II) provides additional opportunities
CC for sustaining the capacity for antibody diversification in cattle
CC essential for immunocompetence by selective breeding strategies that
CC incorporate immunoglobulin gene markers. (III) is unique to cattle and
CC therefore are useful in forensic analysis. The VDJ cassette isolated from
CC cattle provides a natural source of immunoglobulin that is capable of
CC incorporating multiple, both B and T cell epitopes, and is, therefore,
CC most suitable as vaccine vector across species by using species-specific
CC different isotypes, as or if required. The present sequence represents a
CC nucleic acid sequence of plasmid pBf1H1-D3, which is used in the
CC exemplification of the present invention.
XX
SQ Sequence 139 BP; 26 A; 8 C; 43 G; 62 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.0175 Length: 139
Score: 78.00 Matches: 11
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 91.67% Mismatches: 0
Query Match: 92.86% Indels: 0
DB: 10 Gaps: 0

US-09-858-332G-15 (1-12) x ADD08924 (1-139)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
:::|||||
DB 56 TACAACTATACCATTAACCAACCACTATACCATTAAC 21

RESULT 14
ADD08921/c
ID ADD08921 standard; DNA; 205 BP.
XX
AC ADD08921;
XX
DT 01-JAN-2004 (first entry)
XX
DE Bovine germline D nucleic acid sequence D-1 SEQ ID NO:2.
XX
KW bovine; VDJ gene cassette; bovine germline; vaccine; immunotherapy;
KW immunoglobulin; gene; ds.
XX
OS Synthetic.
OS Bos taurus.
OS
OS
PN CA2382766-A1.
XX
PD 20-OCT-2002.
XX
PF 19-APR-2002; 2002CA-02382766.
XX
PR 20-APR-2001; 2001US-0284899P.
XX
PA (SAIN/) SAINI S S.
PA (KAUS/) KAUSHIK A.
XX
PI Saini SS, Kaushik A;
XX
DR WFI; 2003-314127/31.
XX
PT Novel isolated bovine VDJ gene cassette useful for preparing a vaccine

PT vector, and in immunotherapy.
XX
PS Claim 5; Fig 2; 29pp; English.
XX
CC The present invention describes an isolated bovine VDJ gene cassette (I)
CC comprising the 587 nucleotide sequence of ADD08920. Also described: (1) a
CC recombinant plasmid pBf1H1-24 (II); (3) an isolated bovine germline D
CC sequence (III) comprising the sequence of D-1 (ADD08921), D-2 (ADD08922)
CC and/or D-3 (ADD08923); (4) a recombinant plasmid pBf1H1-24; and (5) a
CC recombinant plasmid pBf1H1-D3. (I) or (II) can be used for preparing a
CC vaccine vector. (I) provides the novel ability to develop chimeric
CC immunoglobulin molecule capable of incorporating both linear T cell
CC epitope(s) (CDR1H and CDR2H) as well as conformational B cell epitope(s)
CC (exceptionally long CDR3H). As a result, the germline encodes Bf1H1 VDJ
CC cassette isolated from bovine foetal B cells is suitable for
CC immunoglobulin antigenisation with both B and T cell epitopes and use as
CC vaccine vector. Further multiple epitopes can be incorporated for
CC development of multivalent vaccine by replacing at least a portion of an
CC immunoglobulin molecule with the desired epitope such that the functional
CC ability of both epitope(s) and parent VDJ rearrangement is retained. The
CC antigenised immunoglobulin incorporating both T and B cell epitopes of
CC interest is useful for development of oral vaccines for humans apart from
CC other species including cattle. (III) provides additional opportunities
CC for sustaining the capacity for antibody diversification in cattle
CC essential for immunocompetence by selective breeding strategies that
CC incorporate immunoglobulin gene markers. (III) is unique to cattle and
CC therefore are useful in forensic analysis. The VDJ cassette isolated from
CC cattle provides a natural source of immunoglobulin that is capable of
CC incorporating multiple, both B and T cell epitopes, and is, therefore,
CC most suitable as vaccine vector across species by using species-specific
CC different isotypes, as or if required.
XX
SQ Sequence 205 BP; 42 A; 19 C; 62 G; 82 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.0258 Length: 205
Score: 78.00 Matches: 11
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 91.67% Mismatches: 0
Query Match: 92.86% Indels: 0
DB: 10 Gaps: 0

US-09-858-332G-15 (1-12) x ADD08921 (1-205)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
:::|||||
DB 86 TACAACTATACCATTAACCAACCACTATACCATTAAC 51

RESULT 15
ADD08920/c
ID ADD08920 standard; DNA; 587 BP.
XX
AC ADD08920;
XX
DT 01-JAN-2004 (first entry)
XX
DE Bovine VDJ gene cassette nucleic acid sequence SEQ ID NO:1.
XX
KW bovine; VDJ gene cassette; bovine germline; vaccine; immunotherapy;
KW immunoglobulin; gene; ds.
XX
OS Synthetic.
OS Bos taurus.
OS
OS
PN CA2382766-A1.
XX
PD 20-OCT-2002.
XX
PF 19-APR-2002; 2002CA-02382766.
XX
PR 20-APR-2001; 2001US-0284899P.
XX

Alignment Scores:	
Pred. No.:	0.222
Score:	1731
Percent Similarity:	78.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
OB:	92.86%
	12
	Gaps: 0
	Indels: 0
	Mismatches: 0
	Conservative: 11
	Matches: 1731
	Length: 1731

1 115A6NH15A6NH15A6NH15A6NH15 11

Db 1344 CACAATCATAATCATAATCACAACCAT 1312

Continuation (16 of 57) of ACF67367 from base 1500001 (Photorhabdus luminescens nucleot

WP	Fragment Name	Begin	End
1	1000000000	1000000000	1000000000

WP	ACF67367_01	100001	2100000
WP	ACF67367_01	000001	2100000
WP	ACF67367_01	000001	2100000

WP	ACF67367_03	300001	410000
----	-------------	--------	--------

WP	ACF67367_05	500001	6100000
----	-------------	--------	---------

WP	ACF67367_07	700001	810000
WP	ACF67367_07	700001	810000
WP	ACF67367_07	700001	810000

WP	ACF67367_09	900001	10100000
----	-------------	--------	----------

WP	ACF67367_11	1100001	12100000
----	-------------	---------	----------

WP	ACF67367_13	1300001	14100000
WP	ACF67367_14	1400001	15100000

WP	ACF67367_15	1500001	16100000
WP	ACF67367_16	1600001	17100000

WP	ACF67367_17	1700001	1810000
WP	ACF67367_18	1800001	1820000

WP	ACF67367_19	1900001	2010000
WP	ACE67367_20	2000001	2110000

WP	ACF67367_21	2100001	2210000
WP	ACE67367_33	3300001	3310000

WP ACF67367_23 2300001 2410000

WP	ACF67367_24	2500001	2510000
WP	ACF67367_25	2500001	2610000
WP	ACF67367_26	2600001	2710000
WP	ACF67367_27	2700001	2810000
WP	ACF67367_28	2800001	2910000
WP	ACF67367_29	2900001	3010000
WP	ACF67367_30	3000001	3100000
WP	ACF67367_31	3100001	3210000
WP	ACF67367_32	3200001	3310000
WP	ACF67367_33	3300001	3410000
WP	ACF67367_34	3400001	3510000
WP	ACF67367_35	3500001	3610000
WP	ACF67367_36	3600001	3710000
WP	ACF67367_37	3700001	3810000
WP	ACF67367_38	3800001	3910000
WP	ACF67367_39	3900001	4010000
WP	ACF67367_40	4000001	4110000
WP	ACF67367_41	4100001	4210000
WP	ACF67367_42	4200001	4310000
WP	ACF67367_43	4300001	4410000
WP	ACF67367_44	4400001	4510000
WP	ACF67367_45	4500001	4610000
WP	ACF67367_46	4600001	4710000
WP	ACF67367_47	4700001	4810000
WP	ACF67367_48	4800001	4910000
WP	ACF67367_49	4900001	5010000
WP	ACF67367_50	5000001	5110000
WP	ACF67367_51	5100001	5210000
WP	ACF67367_52	5200001	5310000
WP	ACF67367_53	5300001	5410000
WP	ACF67367_54	5400001	5510000
WP	ACF67367_55	5500001	5610000
WP	ACF67367_56	5600001	5648894

Alignment Scores:	
Pred. No.:	14.6
Score:	1100000
Percent Similarity:	78.00%
Best Local Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	92.86%
DB:	10
Length:	1100000
Matches:	11
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-858-332G-15 (1-12) x ACF67367_15 (1-110000)

QY **1** H1SASNH1SASNH1SASNH1SASNH1S 11

Db 10517 CACAATCACAATCACAATCACAATCAC 10549

RESULT 19

ID ACF65381

AC ACF65381;

DT 20-NOV-2003 (first entry)

DE Photornabhdus luminescens nucleotide sequence #34

KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis

KW antibiotic; biopesticide; virulence factor; disease model; plague;

XX

XX

XXXX

XX

XX
2
g
g
2
2
2
g
g
g
g
g

XX

PA (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Duchaud E, Taourit S, Glaser P, Frangoul L, Kunst F, Danchin A,
 PI Buchreiser C,
 XX WPI; 2003-148459/14.
 XX
 PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX
 PS Claim 1; SEQ ID NO 34; 1205bp; French.
 XX
 CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants.
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens genes
 XX
 SQ Sequence 249878 BP; 75946 A; 56652 C; 49249 G; 66029 T; 0 U; 2 Other;
 XX
 Alignment Scores:
 Pred. No.: 33.5 Length: 249878
 Score: 78.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 92.86% Indels: 0
 DB: 10 Gaps: 0
 US-09-858-332g-15 (1-12) x AC65381 (1-249878)
 QY 1 H1sAenH1sAenH1sAenH1sAenH1sAenH1s 11
 DB 88838 CACCAATCACCAATCACCAATCACCAATCAC 88870
 RESULT 20
 ACNS6255
 ID ACNS6255 standard; cDNA; 537 BP.
 XX
 AC ACNS6255;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Cotton androecium tissue EST Clone ID: LIB3828-033-Q6-N6-D11, SEQ:11036.
 XX
 KW Cotton; plant; EST; expressed sequence tag; transgenic plant; androecium;
 KW variety Nucleon33B; library LIB3828; molecular tag; molecular marker;
 KW genetic mapping; molecular mapping; seed germination; plant growth;
 KW plant quality; plant yield; plant breeding; tissue printing; ss.
 XX
 OS Gossypium hirsutum.
 OS
 XX
 PN US2004123340-A1.
 XX
 PD 24-JUN-2004.
 XX
 PF 12-DEC-2001; 2001US-00021323.

XX
 PR 14-DEC-2000; 2000US-0255619P.
 XX
 PA (DEIK/) DEIKMAN J.
 PA (FENG/) FENG P C C.
 PA (FINC/) FINCHER K L.
 PA (ZIEG/) ZIEGLER T E.
 XX
 PI Deikman J, Feng PCC, Fincher KL, Ziegler TE;
 XX WPI; 2004-479808/45.
 DR
 XX
 PT New isolated nucleic acid molecule that encodes a plant protein or its
 PT fragment, useful for isolating a variety of agronomically significant
 PT genes associated with plant growth, quality or yield, and as molecular
 PT tags to map genes.
 XX
 PS Claim 1; SEQ ID NO 11036; 34pp; English.
 XX
 CC The invention relates to 17880 cotton expressed sequence tags (ESTs;
 CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
 CC from primed or non-primed seeds from variety D950B, mature seeds from
 CC variety Coker 312, Bower 96 Field, and androecium tissue, gynoecium
 CC tissue, developing fibres, carpel walls and septa from variety
 CC Nucleon33B. The invention also relates to substantially purified
 CC proteins or their fragments encoded by nucleic acid molecules of the
 CC invention, and to transformed plants having a nucleic acid construct
 CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
 CC molecular tags to isolate genetic regions, to isolate genes, to map
 CC genes, to determine gene function and to determine whether genes are
 CC members of a particular gene family. The nucleic acid molecules may be
 CC used for isolating a variety of agronomically significant genes
 CC associated with plant growth, quality, yield, and could also serve as
 CC links in metabolic and catabolic pathways. The nucleic acid molecules are
 CC also useful for identifying genes important in initiating and maintaining
 CC seed germination or that may be used to mitigate stresses encountered
 CC during seed germination. The ESTs additionally enable the acquisition of
 CC promoters and cis-regulatory elements which will be useful to express
 CC agronomically significant genes in these tissues and/or other tissues,
 CC and also permits the acquisition of molecular markers useful in breeding
 CC schemes, genetic and molecular mapping, and in cloning of agronomically
 CC significant genes. The nucleic acid molecules are further useful for
 CC detecting the expression level or pattern of a protein or mRNA and for
 CC detecting the presence or quantity of a protein by tissue printing. The
 CC present sequence represents a specifically claimed EST isolated from a
 CC cotton variety Nucleon33B androecium tissue cDNA library (LIB3828). The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the US
 CC patent office at seqdata.uspto.gov/sequence.html?docid=US20040123340
 XX
 SQ Sequence 537 BP; 169 A; 143 C; 60 G; 165 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 0.0918 Length: 537
 Score: 77.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 91.67% Mismatches: 0
 Query Match: 91.67% Indels: 0
 DB: 13 Gaps: 0
 US-09-858-332g-15 (1-12) x ACNS6255 (1-537)
 QY 1 H1sAenH1sAenH1sAenH1sAenH1sAenH1s 12
 DB 240 AATAATCATATATCATATATCATATATCATATAT 275
 RESULT 21
 AAS27823
 ID AAS27823 standard; DNA; 16489 BP.
 XX
 AC AAS27823;
 XX
 DT 07-NOV-2001 (first entry)

XX DE DNA encoding novel signal transduction pathway protein, Seq ID 1483.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ds;
KW acquired immune deficiency syndrome.
XX
XX Homo sapiens.
XX
XX MO200154733-A1.
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001MO-US001312.
PF
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUL-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231422P.
PR 08-SEP-2000; 2000US-0231433P.
PR 08-SEP-2000; 2000US-0231443P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 25-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235483P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240360P.
PR 20-OCT-2000; 2000US-0241212P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.

WPI; 2003-719985/68.

XX New isolated polypeptide useful for diagnosing and treating
PT immunosuppressive conditions such as autoimmune disease and Parkinson's
PT disease.
XX disclosure, SEQ ID NO 1483; 345pp; English.
XX The invention relates to an isolated polypeptide. The polypeptide is
CC useful for diagnosing a pathological condition or a susceptibility to a
CC pathological condition in a subject, by determining the presence or
CC amount of expression of the polypeptide in a biological sample and
CC diagnosing a pathological condition or a susceptibility to a pathological
CC condition based on the presence or amount of expression of the
CC polypeptide. The polypeptide is also useful for identifying a binding
CC partner to the polypeptide, which involves contacting the polypeptide
CC with a binding partner and determining whether the binding partner
CC effects an activity of the polypeptide. The polypeptide or the nucleic
CC acid encoding the polypeptide is useful for preventing, treating, or
CC ameliorating a medical condition, which involves administering the
CC polypeptide or the nucleic acid to a mammalian subject. The nucleic acid
CC is useful for diagnosing a pathological condition or a susceptibility to
CC a pathological condition in a subject, which involves determining the
CC presence or absence of a mutation in the nucleic acid, and diagnosing a
CC pathological condition or susceptibility to a pathological condition
CC based on the presence or absence of the mutation. The polypeptide, the
CC nucleic acid and an antibody to the polypeptide are useful for treating
CC autoimmune disease, Parkinson's disease, silicosis, gastrointestinal
CC disease, atherosclerosis, haemophilia, thrombocytopenia. The polypeptide,
CC the nucleic acid and the antibody are useful as immunosuppressive agents,
CC as adjuvants to enhance immune responses, and as agents to induce higher
CC affinity antibodies and increase serum immunoglobulin concentrations. The
CC present sequence represents DNA encoding a novel human protein. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format direct from USPTO at
CC seqdata.uspto.gov/sequence.html?DocId=20020168711.
XX Sequence 16489 BP; 3783 A; 4714 C; 4435 G; 3557 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 2.9 Length: 16489
Score: 77.00 Matches: 11
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 91.67% Mismatches: 0
Query Match: 91.67% Indels: 0
DB: 10 Gaps: 0

US-09-858-332g-15 (1-12) x ADB94626 (1-16489)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 7017 AATATCATATCATATCATATCATATCATAT 7052

RESULT 23
AADI0057
XX AADI0057 standard; DNA; 1154 BP.
XX
AC AADI0057;
XX
DT 18-SEP-2001 (first entry)
XX
DE llama HCV33 VH region with heavy chain constant regions DNA.
XX
KW Camelid; llama; heavy chain immunoglobulin; pathogen resistance;
KW metabolism modulator; passive immunisation; heavy chain variable domain;
KW VH; anti-R6 antibody; HCV33; azo-dye; ds.
XX
XX llama glama.
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 3..1136
FT /*tag= a

FT /product= "llama HCV33 heavy chain variable domain with
FT heavy chain constant regions"
FT 3..389
FT misc_feature
FT /*tag= b
FT /note= "Corresponds to llama HCV33 heavy chain variable
FT domain"
FT 390..476
FT misc_feature
FT /*tag= c
FT /note= "Corresponds to hinge region DNA"
FT 477..806
FT misc_feature
FT /*tag= d
FT /note= "Corresponds to heavy chain constant region CH2
FT DNA"
FT 807..1133
FT misc_feature
FT /*tag= e
FT /note= "Corresponds to heavy chain constant region CH3
FT DNA"
XX
XX BP118669-A2.
XX
XX 25-JUL-2001.
XX
XX 08-DEC-2000; 2000EB-00310997.
XX
XX 17-DEC-1999; 99EP-00310188.
XX
XX (UNITL) UNILEVER PLC.
XX (UNITL) UNILEVER NV.
XX
XX Frenken LGJ, Van Der Logt CPE, Jobling SA, Tsch Y,
XX WPI; 2001-427157/46.
XX P-P-SDB; AAE05286.
XX
XX Modifying a plant to produce an antibody useful for increasing pathogen
PT resistance or to modulate metabolism comprises introducing a DNA sequence
PT encoding a heavy chain immunoglobulin linked to a peptide that targets a
PT cellular compartment.
XX
XX
XX Example 12; Fig 18; 81pp; English.
XX
XX The present invention relates to a method for modifying a plant to
CC produce an antibody or an active fragment or derivative, or a protein
CC functional equivalent, in a cellular compartment comprising introducing a
CC DNA sequence encoding a heavy chain immunoglobulin, where the DNA is
CC linked to promoters and provided with an additional sequence encoding a
CC peptide capable of targeting heavy chain immunoglobulin to a cellular
CC compartment. The method is used for producing a heavy chain
CC immunoglobulin or an active fragment or derivative, or a protein that is
CC functionally equivalent for increasing the pathogen resistance in a plant
CC or to modulate metabolism in a plant. Under some circumstances it may be
CC desirable to retain the antibody product with the plant rather than
CC extracting and isolating the product. In particular, edible selected
CC antigens may be used in a method of passively immunising an animal,
CC preferably human, against the antigen, e.g., pathogenic organisms. The
CC present DNA sequence encodes a llama (camelid) anti-R6 (R6 is an azo-
CC dye) antibody (designated as HCV33) heavy chain variable domain (VH)
CC linked to the non-classical heavy chain constant regions denoted as hinge
CC -CH2-CH3
XX
XX Sequence 1154 BP; 293 A; 341 C; 321 G; 199 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 0.267 Length: 1154
Score: 76.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 90.48% Indels: 0
DB: 5 Gaps: 0

US-09-858-332g-15 (1-12) x AADI0057 (1-1154)

QY 2 AsnHisAsnHisAsnHisAsnHisAsn 12

CC Aspergillus fumigatus and Cryptococcus neoformans. The methods are useful
CC to identify agents that may be used in the treatment of fungal
CC infections. AAS23381-AAS23442 represent C. albicans essential genes
XX

SO Sequence 1521 BP, 536 A; 289 C; 321 G; 375 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.353	Length:	1521
Score:	76.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	90.48%	Indels:	0
DB:	4	Gaps:	0

US-09-858-332G-15 (1-12) x AAS23387 (1-1521)

QY 2 AsnHisAsnHisAsnHisAsnHisAsn 12
Db 173 AACCAACACCAACCAACCAACCAACCAAC 205

RESULT 26

AB231758
ID AB231758 standard; DNA; 1521 BP.

AC AB231758;

DT 30-JAN-2003 (first entry)

DE Candida albicans essential gene SEQ ID NO 6045.

XX Fungus; yeast; tetracycline; promoter; GRACE strain; biosynthesis;
KW signal transduction; DNA replication; cell division; growth;
KW proliferation; Candida albicans; fungicide; antifungal; gene; ss.

XX

OS Candida albicans.

XX WO200253728-A2.

XX 11-JUL-2002.

XX 26-DEC-2001; 2001WO-US049486.

XX 29-DEC-2000; 2000US-0259128P.

XX 20-FEB-2001; 2001US-00792024.

XX 22-AUG-2001; 2001US-0314050P.

XX (ELIT-) ELITRA PHARM INC.

XX Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
DR WPI: 2002-566694/60.
DR P-PSDB: ABP73208.

XX

PT Constructing strains for identifying gene products as effective targets
PT for therapeutic intervention, by inactivating in the strain one allele of
PT a gene and placing other allele of the gene under conditional expression.

XX

PS Claim 37; SEQ ID NO 6045; 167pp + Sequence Listing; English.

XX The invention relates to constructing (M1) a strain of diploid fungal
CC cells in which both alleles of a gene are modified, comprising modifying
CC one allele by insertion or replacement by a cassette having an
CC expressible selectable marker and modifying other allele by
CC recombination, of a promoter replacement fragment with a heterologous
CC promoter, so that expression of the second allele is regulated by the
CC promoter. (M1) is useful for constructing a strain of diploid fungal
CC cells in which both alleles of a gene are modified. The diploid fungal
CC cells having both alleles modified are useful for identifying a gene that
CC is essential to the survival and/or pathogenicity of a fungus, a gene that
CC contributes to the virulence and/or pathogenicity of a fungus, a gene
CC that contributes to the resistance of a diploid fungus to an antifungal
CC agent, an antifungal agent that inhibits the growth of a diploid fungus
CC and for identifying a therapeutic agent for treatment of a mammalian

CC disease. (M1) is useful for identifying a compound which modulates the
CC activity of a gene product, preferably enzymatic activity, carbon
CC compound catabolism, biosynthetic, transporter, transcriptional,
CC translational, signal transduction, DNA replication and cell division
CC activity. The method is useful for identifying a compound having the
CC ability to inhibit growth or proliferation of C. albicans cells and for
CC treating infection by C. albicans. The present sequence is that of an
CC essential Candida albicans gene used in the method of the invention.

CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office
XX

SO Sequence 1521 BP, 536 A; 289 C; 321 G; 375 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.353	Length:	1521
Score:	76.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	90.48%	Indels:	0
DB:	6	Gaps:	0

US-09-858-332G-15 (1-12) x AB231758 (1-1521)

QY 2 AsnHisAsnHisAsnHisAsnHisAsn 12
Db 173 AACCAACACCAACCAACCAACCAACCAAC 205

RESULT 27

AA165447/C
ID AA165447 standard; DNA; 4565 BP.

XX AA165447;

XX 10-DEC-2001 (first entry)

XX Nucleotide sequence of a maize activator element (Ac).

XX Activator element; Ac; transposable element; Ds element; transposon;
KW gamete-specific promoter; suicide gene; gametophytic suicide trait; GST;
KW ss.

XX Zea mays.

XX WO200164926-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006249.

XX 28-FEB-2000; 2000US-0185524P.

XX (UYVA) UNIV YALE.

XX Dellaporta SL, Moreno MA;

XX WPI: 2001-602568/68.

XX

PT Genetic construct for the control of transgenes in transgenic plants
PT comprises a sex-specific promoter operatively linked to a suicide gene
PT that selects against male or female gametes containing the suicide gene.

XX Example 1; Page 73-74; 81pp; English.

XX The present sequence represents a maize Activator element (Ac), which is
CC a transposable element. 5' and 3' fragments of the present sequence were
CC combined by cloning to produce a synthetic Ds element, which was used to
CC produce the construct of the invention. The specification describes a
CC nucleic acid construct, comprising a male gamete- or female gamete-
CC specific promoter operably linked to a suicide gene, both of which are
CC linked to a gene of interest, a transposable element, a transposon and
CC the encoding nucleic acid. The construct controls the unwanted spread of
CC heterologous traits in plants. It also provides genetic systems which can

CC be used for the elimination of a gametophytic suicide trait (GST) and for
CC the selection of unlinked transpositions. The construct is applicable for
CC any commercially grown plant, including fruit, seed, oil, protein or hay
CC production, animal grazing, golf courses, lawns, erosion control,
CC landscaping, green manure, producing food additives, pulp and wood
CC production and smoking products

SO Sequence 4565 BP; 1344 A; 994 C; 978 G; 1249 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.07	Length:	4565
Score:	76.00	Matches:	11
Percent Similarity:	91.67%	Conservative:	0
Best Local Similarity:	91.67%	Mismatches:	1
Query Match:	90.48%	Indels:	0
DB:	4	Gaps:	0

US-09-858-332g-15 (1-12) x AAT65447 (1-4565)

QY 1 HisAgnHisAgnHisAgnHisAgnHisAgn 12
|||||
Db 3146 CACAAACCAACCAACCAACCAACCAACCAAC 3111

RESULT 28
ABX93542
ID ABX93542 standard; DNA; 4565 BP.

XX ABX93542;
XX

DT 06-JUN-2003 (first entry)

XX Transposable element associated nucleic acid.

XX ds; transposon modification; gene repression; antibiotic resistance;
KW human gamete; transposon; shoot; heart disease; phenylketonuria;
KW Alzheimer's disease; aging; emphysema; cancer; osteoporosis; diabetes;
KW sleep apnea; sickle cell anaemia; stroke; multiple sclerosis;
KW muscular dystrophy; buritis; diverticular disease; glaucoma; blindness;
KW cataract; attention deficit disorder; dyslexia; hypertension;
KW schizophrenia; depression; bulimia nervosa; migraine; anorexia nervosa;
KW heart attack; allergy; baldness; metabolic disorder; blood clotting;
KW haemophilia; thalassemia; neuromuscular disease; erectile dysfunction;
KW pregnancy associated disease; infertility; gastritis; obesity;
KW immune system disorder; blood cell surface antigen recognition disease;
KW nondisjunction; epilepsy; spinal cord disorder; pituitary gland disorder;
KW Parkinson's disease; osteoarthritis; cystic fibrosis; arterial disease;
KW mania; cardiomyopathy.

XX Unidentified.

OS US2002199216-A1.

XX 26-DEC-2002.

PF 01-MAY-2002; 2002US-00138221.

PR 01-MAY-2001; 2001US-0287882P.

XX (MACR/) MACRAE A F.

XX Macrae AF;

XX WPI; 2003-329235/31.

PT Modifying transcription within an organism or repressing transcription of
PT targeted gene within a cell, by introducing transposase-encoding effector
PT molecule into organism or introducing transposase element into cell.

XX Disclosure; Page 34-37; 46pp; English.

CC The invention relates to a method of modifying (increasing, decreasing or
CC altering) transcription within an organism or repressing transcription of
CC at least one targeted gene within a cell comprising introducing a

CC transposase-encoding effector molecule into the organism or introducing a
CC transposase element into the cell. The method is useful for modifying a
CC (increasing, decreasing or altering) transcription within an organism or
CC for repressing transcription of at least one targeted gene within a cell.
CC The method is useful for determining the functions of unknown function
CC genes within an organism. The method is useful for repressing
CC transcription within an organism e.g. transcription of a gene encoding
CC antibiotic resistance, for repressing targeted genes within human gametes
CC (sperm and egg). The method is useful for repressing transcription within
CC an in vitro cell system. The method is useful for repressing transcription
CC during the cell cycle. The method is also useful for
CC repressing transcription within a transplanted organ, shoot, or body
CC part. The method is also useful for modifying the expression of genes
CC associated with heart disease, phenylketonuria, Alzheimer's disease,
CC aging, cancer, osteoporosis, diabetes, sleep apnea, sickle cell anaemia,
CC multiple sclerosis, muscular dystrophy, buritis, emphysema, diverticular
CC disease, glaucoma, blindness, cataracts, attention deficit disorder,
CC dyslexia, hypertension, schizophrenia, mania, depression, bulimia
CC nervosa, anorexia nervosa, stroke, heart attack, allergies, migraine,
CC baldness, metabolic disorders, blood clotting, haemophilia,
CC thalassemia, neuromuscular diseases, diseases of pregnancy, erectile
CC dysfunction, infertility, immune system disorders, epilepsy, obesity,
CC antigen recognition disease, nondisjunction, pituitary gland disorder,
CC gastroenteritis, spinal cord disorders, cystic fibrosis, arterial disease
CC and cardiomyopathy. The present sequence represents the transposable
CC element associated nucleic acid. Note: the present sequence is shown in
CC the appendix of the specification but no reference is made to it
XX elsewhere in the specification

SQ Sequence 4565 BP; 1249 A; 978 C; 994 G; 1344 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.07	Length:	4565
Score:	76.00	Matches:	11
Percent Similarity:	91.67%	Conservative:	0
Best Local Similarity:	91.67%	Mismatches:	1
Query Match:	90.48%	Indels:	0
DB:	8	Gaps:	0

US-09-858-332g-15 (1-12) x ABX93542 (1-4565)

QY 1 HisAgnHisAgnHisAgnHisAgnHisAgn 12
|||||
Db 1420 CACAAACCAACCAACCAACCAACCAACCAAC 1455

RESULT 29
ABX93541/C

ID ABX93541 standard; DNA; 4810 BP.

XX ABX93541;
XX

DT 06-JUN-2003 (first entry)

XX DNA encoding the maize transposable element activator (Ac9).

XX Maize; ds; transposable element activator; transcription modification;
KW gene repression; antibiotic resistance; human gamete; transposon; shoot;
KW heart disease; phenylketonuria; Alzheimer's disease; aging; emphysema;
KW cancer; osteoporosis; diabetes; sleep apnea; sickle cell anaemia; stroke;
KW glaucoma; blindness; muscular dystrophy; buritis; diverticular disease;
KW hypertension; schizophrenia; depression; attention deficit disorder; dyslexia;
KW anorexia nervosa; heart attack; allergy; baldness; metabolic disorder;
KW blood clotting; haemophilia; thalassemia; neuromuscular disease; Ac9;
KW pregnancy associated disease; erectile dysfunction; infertility; obesity;
KW immune system disorder; blood cell surface antigen recognition disease;
KW nondisjunction; epilepsy; gastritis; spinal cord disorder;
KW pituitary gland disorder; Parkinson's disease; osteoarthritis;
KW cystic fibrosis; arterial disease; mania; cardiomyopathy; waxy locus.
XX Zea mays.

PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140921P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143452P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0143624P.
PR 16-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144864P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145066P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148654P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.

PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152232P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 0 485
Score: 73.00
Percent Similarity: 100.00%
Best Local Similarity: 90.91%
Query Match: 86.90%
DB: 3
Gaps: 0

US-09-858-332g-15 (1-12) x AAC6337 (1-862)

OY 1 H1sAenH1sAenH1sAenH1sAenH1s 11
Db 462 CATATCATATCATATCATATCATATCAT 430

RESULT 31

AAC46056/c

ID AAC46056 standard; DNA; 862 BP.

XX AAC46056;

AC AAC46056;

XX 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 48749.
XX
XX Hybridisation assay; Genetic mapping; gene expression control;
KM protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132408P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132853P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137232P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-014187P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145813P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0149931P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.

PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152353P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154779P.
PR 20-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155658P.
PR 26-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161355P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.: 0.485
Score: 73.00
Percent Similarity: 100.00%
Best Local Similarity: 90.91%
Query Match: 86.90%
DB: 3
Matches: 862
Conservative: 10
Mismatch: 1
Indels: 0
Gaps: 0

US-09-858-332g-15 (1-12) x AAC46056 (1-862)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHis 11

Db 462 CATATCATATCATATCATATCATATCATATCATAT 430

RESULT 32

AAC52042/c

AAC52042 standard; DNA; 960 BP.

AC AAC52042;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 70262.

KM Hybridisation assay; generic mapping; gene expression control;
KM protein identification; signal transduction pathway; metabolic pathway;
KM promoter; termination sequence; ss.

OS Arabidopsis thaliana.

EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130443P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 04-MAY-1999; 99US-0133407P.

PR 05-MAY-1999; 99US-0133484P.

PR 06-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132487P.

PR 11-MAY-1999; 99US-0132863P.

PR 14-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134788P.

PR 19-MAY-1999; 99US-0134788P.

PR 20-MAY-1999; 99US-0134941P.

PR 21-MAY-1999; 99US-0135124P.

PR 24-MAY-1999; 99US-0135353P.

PR 25-MAY-1999; 99US-0135629P.

PR 27-MAY-1999; 99US-0136021P.

PR 28-MAY-1999; 99US-0136392P.

PR 01-JUN-1999; 99US-0136782P.

PR 03-JUN-1999; 99US-0137222P.

PR 04-JUN-1999; 99US-0137528P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 14-JUN-1999; 99US-0138847P.

PR 16-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 17-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139452P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 21-JUN-1999; 99US-0139763P.

PR 22-JUN-1999; 99US-0139817P.

PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-015088P.
PR 22-JUL-1999; 99US-015085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147483P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149417P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.

PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155658P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 16-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.: 0.541 Length: 960
Score: 73.00 Matches: 10
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 90.91% Mismatches: 0
Query Match: 86.90% Indels: 0
DB: 3 Gaps: 0

US-09-858-332g-15 (1-12) x AAC52042 (1-960)

Qy 1 HisAsnHisAsnHisAsnHisAsnHis 11

Db 564 CATATCATATATCATATCATATATCAT 532

RESULT 33

ABQ76560 standard; cDNA; 3042 BP.

ID ABQ76560;

AC ABQ76560;

XX 21-NOV-2002 (first entry)

DE C. albicans BAX-associated cDNA fragment SEQ ID 545.
XX Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;

KW antineumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
ds.
XX
XX Homo sapiens.
OS
XX
XX MO200200928-A2.
XX
XX
XX
XX 03-JUN-2002.
PD
XX
XX 02-JUL-2001; 2001WO-EP007537.
PF
XX
XX 30-JUN-2000; 2000DE-01032529.
PR
XX 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIC-) EPIDENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX
XX WPI; 2002-130909/17.
DR
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
XX
XX Claim 1; SEQ ID NO 1603; 32pp + Sequence Listing; German.
PS
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
CC
SQ Sequence 6236 BP; 1410 A; 206 C; 1688 G; 2932 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 4.81 Length: 6236
Score: 72.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 85.71% Indels: 0
DB: Gaps: 0
US-09-858-332g-15 (1-12) x ABU33630 (1-6236)
QY
XX 1 H18A8NH18A8NH18A8NH18A8NH18A8NH18 11
ID A8N80202/c
DB 4241 CATACCAACACATACCATACCAATACCAAC 4209
RESULT 36
ID A8N80202 standard; DNA: 6236 BP.
XX
XX A8N80202;
AC
XX
XX 15-JUL-2002 (first entry)
DT
XX
XX Human chemically modified disease associated gene SEQ ID NO 219.
DE
XX
XX Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
KW heart disease; epilepsy; histone deacetylation; muscular dystrophy;
KW dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
KW antidiabetic; cytosine; anticonvulsant; ds.
XX
XX Homo sapiens.
OS
XX Synthetic.
XX
XX MO200200927-A2.

XX
XX 03-JUN-2002.
PD
XX
XX 02-JUL-2001; 2001WO-EP007536.
PF
XX
XX 30-JUN-2000; 2000DE-01032529.
PR
XX 01-SEP-2000; 2000DE-01043826.
XX
XX
XX (EPIC-) EPIDENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX
XX WPI; 2002-130908/17.
DR
XX
XX Novel nucleic acid useful for diagnosis and therapy of diseases
PT associated with development genes such as diabetes; comprises a sequence
PT of a segment of chemically pretreated DNA of genes associated with
PT development.
XX
XX
XX Claim 1; SEQ ID NO 219; 27pp; English.
PS
XX
XX The invention relates to a nucleic acid (I) comprising a sequence at
CC least 18 bases in length of a segment of chemically pretreated DNA (II)
CC of genes associated with development selected from 87 genes listed in the
CC specification such as ACCPN, ADFN, or APFI and comprising one of 350
CC sequences (ABN79984-ABN80333) or their complements. The invention is
CC useful for the diagnosis or therapy of diseases associated with
CC development genes, in particular disease related to homeobox containing
CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
CC associated with congenital heart disease, epilepsy, diseases related to
CC histone deacetylation, Curraño syndrome, diseases related with the
CC development of the brain and limb girdle muscular dystrophy and dwarfism.
CC Oligomers specific to each of the genes are useful for detecting the
CC methylation state of all CpG dinucleotides within the 350 sequences or
CC (II) and their complementary sequences, as primer oligonucleotides for
CC the amplification of the 350 sequences, (II) and/or their complements and
CC as oligomer probes for detecting the cytosine methylation state and/or
CC single nucleotide polymorphisms (SNPs). Note: The sequence data for this
CC patent did not form part of the printed specification but is based on
CC sequence information supplied to Derwent by the European Patent Office
XX
SQ Sequence 6236 BP; 1410 A; 206 C; 1688 G; 2932 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 4.81 Length: 6236
Score: 72.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 85.71% Indels: 0
DB: Gaps: 0
US-09-858-332g-15 (1-12) x A8N80202 (1-6236)
QY
XX 1 H18A8NH18A8NH18A8NH18A8NH18A8NH18 11
ID A8K94412
DB 4241 CATACCAACACATACCATACCAATACCAAC 4209
RESULT 37
ID A8K94412 standard; DNA: 74037 BP.
XX
XX A8K94412;
AC
XX
XX 27-AUG-2002 (first entry)
DT
XX
XX DNA encoding endothelin converting enzyme 1 (ECE-1) #2.
DE
XX
XX Endothelin; EDN; endothelin converting enzyme 1; ECE 1; EDNR;
KW endothelin receptor; signaling system; cardiovascular disease;
KW coronary heart disease; hypertension; arteriosclerosis; angiogenesis;
KW fatty acid metabolism; diabetes; familial hypercholesterolaemia;
KW forensic marker; transgenic animal; solid support; SNP;
KW cardiovascular regulator; gene; ds; single nucleotide polymorphism.

XX	Homo sapiens.
OS	
FH	Key
FT	Variation
FT	/tag= a
FT	/standard_name= "Single nucleotide polymorphism"
FT	replacement(8625,T)
FT	*tag= b
FT	/standard_name= "single nucleotide polymorphism"
FT	replacement(9351,G)
FT	*tag= c
FT	/standard_name= "Single nucleotide polymorphism"
FT	replacement(9439,C)
FT	*tag= d
FT	/standard_name= "Single nucleotide polymorphism"
XX	
FN	W0200224747-A2.
PX	
PD	28-MAR-2002.
XX	
PF	31-AUG-2001; 2001WO-EP010087.
PX	
PR	19-SEP-2000; 2000EP-00120123.
PA	(EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
PI	
PJ	Brinkmann U, Hoffmeyer S;
DR	
XX	WPL; 2002-435060/46.
PT	
PT	Novel polynucleotide of the endothelin/endothelin converting enzyme signaling system associated with cardiovascular disease, useful for treating the disease.
PT	
XX	
XX	Claim 1; Page; 190pp; English.
XX	The invention describes a polynucleotide (I) of the endothelin (EDN)/endothelin converting enzyme (ECE)/receptors of EDN and ECE (EDNR) signaling system which is associated with a cardiovascular disease. (I), the gene encoding EDN, ECE or EDNR (II) or a vector (III) expressing (I) or (II) is useful for producing cells capable of expressing a molecular variant polypeptide which is associated with a cardiovascular disease. (II), (III), the EDN, ECE or EDNR polypeptide, or a cell expressing a molecular variant gene comprising (I) is useful for identifying and obtaining a pro-drug or drug capable of modulating the activity of a molecular variant of a polypeptide of the EDN/EDNR/ECE signaling system or its gene product, or for identifying and obtaining an inhibitor of the activity of a molecular variant of a polypeptide of the EDN/EDNR/ECE signaling system or its gene product. The isolated proteins and polynucleotides encoding them are useful for preparation of a pharmaceutical composition for treating a cardiovascular disease such as coronary heart disease, hypertension, atherosclerosis, or related to hypercholesterolaemia or fatty acid metabolism e.g. diabetes and familial EDN/BCE/EDNR signalling animal and in creation of a solid support creating a transgenic animal and genes, vectors, polypeptides, antibodies or compounds polynucleotides, genes, vectors, polypeptides, antibodies or host cells of the invention. This sequence encodes a fragment of the cardiovascular regulator Endothelin converting enzyme 1 (ECE-1). Note: This sequence does not appear in the specification but has been obtained from GenBank using information given in the invention
CC	
CC	
SQ	Sequence 74037 BP; 18330 A; 19332 C; 19180 G; 17195 T; 0 U; 0 Other;
Alignment Scores:	
Align. No.:	78.5 Length: 74037
Score:	71.00 Matches: 10
Percent Similarity:	100.00% Conservative: 1
Best Local Similarity:	90.91% Mismatch: 0
Query Match:	84.52% Indels: 0
Gaps:	0

```

US-09-858-332G-15 (1-12) x ABE94412 (1-74037)
OY      1 HisaEnHisaEnHisaEnHisaEnHis 11
Db       37539 AAATAATCATTAATCATTAATCATTAATCATTAATCAT 37571

RESULT 38
AAA26665
ID AAA26665 standard; cDNA; 904 BP.
XX
AC AAA26665;
XX
DT 23-JUN-2000 (first entry)
XX
DE Candida albicans polynucleotide sequence #3.
KM Candida albicans infection; growth; survival; medication; AIDS;
KW vulvovaginitis; immunocompromised patient; treat; ss.
XX
OS Candida albicans.
XX
PN EP82401-A2.
XX
PD 01-MAR-2000.
PF 23-DEC-1998; 98EP-00310694.
PR 14-AUG-1998; 98GB-00017796.
PA (JANC ) JANSSEN PHARM NV.
XX
PI Contreras RH, Nelissen B, De Backer MD, Luyten WHML, Viaene JE;
PI Logghe MG;
XX
XX WPI; 2000-258614/23.
PT Essential polypeptides isolated from Candida albicans, useful in the
PT treatment of diseases caused by C.albicans, especially in
PT immunocompromized subjects, e.g., AIDS patients.
XX
PS Claim 1; Page 12; 133pp; English.
XX
XX This sequence represents a polynucleotide sequence encoding a polypeptide
CC that is critical for the survival and growth of Candida albicans. The C.
CC albicans nucleic acid molecules of the invention may be used as probes
CC and primers for detecting homologous nucleic acid molecule sequences. The
CC polypeptides and nucleic acid molecules and compounds identified as
CC selectively modulating the expression of the polypeptides, may be used as
CC medications or for the preparation of a medicament to treat C.albicans
CC associated diseases, especially in AIDS patients and to treat
CC vulvovaginitis in otherwise healthy females. The use of the polypeptides
CC and polynucleotide sequences to treat C.albicans associated diseases has
CC fewer side effects and less toxicity than previously used methods such as
CC the use of amphotericin. This method is therefore especially suitable for
XX immunocompromised patients, such as AIDS patients
XX

SQ Sequence 904 BP; 340 A; 140 C; 151 G; 273 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:          1,24           Length:         904
Score:              70.00          Matches:        10
Percent Similarity: 83.33%         Conservative:   0
Best Local Similarity: 83.33%       Mismatches:    2
Query Match:        83.33%         Indels:        0
DBI:                3             Gaps:          0

US-09-858-332G-15 (1-12) x AAA26665 (1-904)
OY      1 HiseEnHiseEnHiseEnHiseEnHiseEn 12
Db       24 CAACAACAATCAATGCATTAACAACAATCAAT 59
```

```

RESULT 39
AAV19873/c
ID AAV19873 standard; DNA; 2556 BP.
XX
XX AAV19873;
AC
XX
XX 23-JUN-1998 (first entry)
DE Drosophila p70S6K gene.
XX
XX p70S6K; p70 S6 kinase; enzyme; kinase signalling pathway; ss.
XX
XX Drosophila melanogaster.
OS
XX
XX Key Location/Qualifiers
FH CDS 474..2387
FT /*tag= a
FT
XX
XX MO9803662-A2.
PN
XX
XX 29-JAN-1998.
PD
XX
XX 11-JUL-1997; 97MO-EP003680.
PR
XX
XX 24-JUL-1996; 96GB-00015498.
PA
XX (NOVS ) NOVARTIS AG.
PI
XX Stewart M, Kozma S, Thomas G;
XX
XX WPI; 1998-120783/11.
DR
XX P-PSDB; AAM52294.
PT
XX New isolated Drosophila p70 S6 kinase - is used to develop products for
PT studying the kinase signalling pathway and for modulating the kinase
XX activity.
XX
XX Claim 3; Page 38-42; 47pp; English.
XX
XX This sequence encodes the Drosophila p70 S6 kinase (p70S6) of the
XX invention. The products are used in the study of the p70 S6 kinase
XX enzyme. They can be used for defining the p70 S6 kinase signalling
XX pathway and to develop agents for modulating the kinase activity. The
XX products can also be used for the production of antibodies and for
XX detection
XX
XX SQ Sequence 2556 BP; 717 A; 663 C; 643 G; 533 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 3.54 Length: 2556
XX Score: 70.00 Matches: 10
XX Percent Similarity: 91.67% Conservative: 1
XX Best Local Similarity: 83.33% Mismatches: 1
XX Query Match: 83.33% Indels: 0
XX DB: 2 Gaps: 0
XX
XX US-09-858-332G-15 (1-12) x AAV19873 (1-2556)
QY 1 H1SA8NH1SA8NH1SA8NH1SA8NH1SA8NH1SA8N 12
DB 2325 CATTAATCATTAATCATTAATCATTAATCATTAAT 2290
XX
XX RESULT 40
XX AAA61847
XX ID AAA61847 standard; DNA; 5511 BP.
XX
XX AAA61847;
AC
XX
XX 28-OCT-2000 (first entry)
DE
XX Cryptosporidium parvum Iowa isolate GP900 ORF.
XX
XX GP900; Iowa isolate; glycoprotein; antibody; cryptosporidiosis;
XX

```

```

KM competitive inhibition; attachment; invasion; ligand binding; sporozoite;
KM merozoite; diarrhoea; protozoacide; open reading frame; ORF; ds.
XX
XX Cryptosporidium parvum.
OS
XX
XX Key Location/Qualifiers
FH CDS 1..5511
FT /*tag= a
FT /product= "Cryptosporidium parvum Iowa isolate GP900"
FT /note= "No stop codon given"
XX
XX US6071518-A.
PN
XX
XX 06-JUN-2000.
PD
XX
XX 12-SEP-1997; 97US-00928361.
PR
XX
XX 29-MAY-1992; 92US-00891301.
PR
XX 01-JUN-1993; 93US-00071880.
PR
XX 03-APR-1995; 95US-00415751.
PR
XX 14-AUG-1996; 96US-00700651.
PR
XX 13-SEP-1996; 96US-0026062P.
PA
XX (REGC ) UNIV CALIFORNIA.
PI
XX
XX Petersen C;
XX
XX WPI; 2000-422065/36.
DR
XX P-PSDB; AAB11726.
PT
XX New GP900 protein fragments and fusion proteins of Cryptosporidium
PT parvum, useful for detecting the presence of the parasite, and diagnosing
PT or treating Cryptosporidium infections by competitive inhibition of the
PT function of GP900.
XX
XX Claim 16; Col 37-42; 59pp; English.
XX
XX The invention relates to the GP900 glycoprotein of the protozoan
XX Cryptosporidium parvum, DNA encoding it, GP900 fragments, and fusion
XX proteins comprising GP900 fragments. The invention also relates to the
XX administration of GP900 or fragments thereof to a host to elicit anti-
XX GP900 antibody production, and to a method of cryptosporidiosis treatment
XX or prophylaxis comprising administration of anti-GP900 antibodies to an
XX individual. Cryptosporidium parvum GP900 and GP900 fragments are able to
XX competitively inhibit sporozoite or merozoite attachment or invasion, and
XX are also useful for the generation of anti-GP900 antibodies. The
XX antibodies also inhibit sporozoite or merozoite attachment/invasion and
XX additionally inhibit the binding of GP900 ligands to GP900. GP900
XX proteins, fragments and antibodies may therefore be used to treat or
XX prevent cryptosporidiosis. Infection with Cryptosporidium is a common
XX cause of diarrhoea in humans and causes life-threatening diarrhoea in
XX immunocompromised persons. Cryptosporidiosis can be contracted from
XX contaminated municipal water supplies (e.g., public swimming pools). It
XX is also a cause of disease in animals, resulting in financial losses in
XX agriculture. GP900 fragments, fusion proteins and antibodies may also be
XX used for the diagnosis of Cryptosporidium parvum infections, and for the
XX detection of the parasite in the environment. The present sequence
XX represents the open reading frame (ORF) encoding the GP900 protein of the
XX Iowa isolate of Cryptosporidium parvum
XX
XX SQ Sequence 5511 BP; 1964 A; 1204 C; 944 G; 1399 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 7.69 Length: 5511
XX Score: 70.00 Matches: 10
XX Percent Similarity: 91.67% Conservative: 1
XX Best Local Similarity: 83.33% Mismatches: 1
XX Query Match: 83.33% Indels: 0
XX DB: 3 Gaps: 0
XX
XX US-09-858-332G-15 (1-12) x AAA61847 (1-5511)
QY 1 H1SA8NH1SA8NH1SA8NH1SA8NH1SA8NH1SA8N 12

```

```

Db      1443  TACAAACCAACCAACCAACCAACCGCAACCAAC 1478
          ::::::::::::::::::::::::::::::::::::::
RESULT 41
ABT04776
ID      ABT04776 standard; DNA; 5511 BP.
XX
XX      ABT04776;
AC
XX      27-SEP-2002 (first entry)
DT
XX      C parvum GP900 gene fragment SEQ ID NO: 2.
DE
XX      Cryptosporidium detection; GP900; P68; cryptopain; cryptosporidiosis;
XX      gene; ds.
OS
XX      Cryptosporidium parvum.
PN      NO200194631-A1.
PD
XX      13-DEC-2001.
XX
XX      14-MAY-2001; 2001WO-US015624.
PF
XX      06-JUN-2000; 2000US-00588995.
PR
XX      (REGC ) UNIV CALIFORNIA.
PA
XX      Petersen C, Barnes DA, Nelson RG, Gut J,
PI      WPI; 2002-566447/60.
XX
XX      Detecting Cryptosporidium in biological and environmental samples and
PT      diagnosis of cryptosporidiosis involves, contacting the sample with
PT      Cryptosporidium GP900, P68 or cryptopain antigen, antibody, DNA or RNA.
PS
XX      Disclosure; Page 99-101; 157pp; English.
XX
XX      The present invention relates to a method of detecting Cryptosporidium in
CC      biological and environmental samples, and of diagnosing
CC      cryptosporidiosis. This involves obtaining a sample and contacting it
CC      with Cryptosporidium GP900, P68 or cryptopain antigen, antibody, DNA or
CC      RNA, or its variant, mutant or fragment. The method is also useful for
CC      detecting and identifying individual Cryptosporidium isolates based on
CC      the genetic characteristics, and for diagnosis of prior or concurrent
CC      Cryptosporidium infection. The present sequence is a C. parvum coding
CC      sequence used in the exemplification of the invention
CC
XX      Sequence 5511 BP; 1964 A; 1204 C; 944 G; 1399 T; 0 U; 0 Other;
SQ
XX
XX      Alignment Scores:
XX      Pred. No.: 7.69
XX      Score: 70.00
XX      Percent Similarity: 91.67%
XX      Best Local Similarity: 83.33%
XX      Query Match: 83.33%
XX      DB: 6
XX      Gaps: 0
XX
XX      US-09-858-332G-15 (1-12) x ABT04776 (1-5511)
OY      1 HisaenHisaenHisaenHisaenHisaenHis 12
Db      1443  TACAAACCAACCAACCAACCAACCGCAACCAAC 1478
          ::::::::::::::::::::::::::::::::::::::
RESULT 42
AAA61846
ID      AAA61846 standard; DNA; 7334 BP.
XX
XX      AAA61846;
AC
XX      28-OCT-2000 (first entry)
DT
XX      Cryptosporidium parvum Iowa isolate GP900 DNA.
DE

```

[illegible]

Query Match: 83.33% Indels: 0
DB: 3 Gaps: 0

US-09-858-332g-15 (1-12) x AAA61846 (1-7334)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
:::|||||
DB 3111 TACACACCAACCAACCAACCAACCAACCAAC 3146

RESULT 43
AB074775
ID AB074775 standard; DNA; 7334 BP.

AC AB074775;
XX
XX
DT 27-SEP-2002 (first entry)
XX
DE C parvum GP900 gene fragment SEQ ID NO: 1.
XX
XX
KM CRYPTOSPORIDIUM detection; GP900; P68; cryptopain; cryptosporidiosis;
KM gene; ds.
XX
OS CRYPTOSPORIDIUM parvum.
XX
PN WO200194631-A1.
XX
PD 13-DEC-2001.
XX
PF 14-MAY-2001; 2001WO-US015624.
XX
XX
PR 06-JUN-2000; 2000US-00588995.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Petersen C, Barnes DA, Nelson RG, Gut J;
XX
XX
XX
DR WPI; 2002-566447/60.
XX
XX
PT Detecting Cryptosporidium in biological and environmental samples and
PT diagnosis of cryptosporidiosis involves, contacting the sample with
PT CRYPTOSPORIDIUM GP900, P68 or cryptopain antigen, antibody, DNA or RNA.
XX
XX
PS Disclosure; Page 97-99; 157pp; English.
XX
XX
CC The present invention relates to a method of detecting Cryptosporidium in
CC biological and environmental samples, and of diagnosing
CC cryptosporidiosis. This involves obtaining a sample and contacting it
CC with CRYPTOSPORIDIUM GP900, P68 or cryptopain antigen, antibody, DNA or
CC RNA, or its variant, mutant or fragment. The method is also useful for
CC detecting and identifying individual Cryptosporidium isolates based on
CC the genetic characteristics, and for diagnosis of prior or concurrent
CC CRYPTOSPORIDIUM infection. The present sequence is a C. parvum coding
CC sequence used in the exemplification of the invention
XX
XX
SQ Sequence 7334 BP; 2627 A; 1536 C; 1263 G; 1908 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 10.3 Length: 7334
Score: 70.00 Matches: 10
Percent Similarity: 91.67% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 83.33% Indels: 0
DB: 6 Gaps: 0

US-09-858-332g-15 (1-12) x AB074775 (1-7334)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
:::|||||
DB 3111 TACACACCAACCAACCAACCAACCAACCAAC 3146

RESULT 44
AB076608
ID AB076608 standard; cDNA; 24076 BP.

XX
AC AB076608;
XX
DT 21-NOV-2002 (first entry)
XX
XX
DE C. albicans BAX-associated cDNA fragment SEQ ID 641.
XX
XX
KM Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
KM vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
KM apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
KM neurodegeneration; cell death; ss.
XX
XX
OS Candida albicans.
XX
PN WO200264766-A2.
XX
PD 22-AUG-2002.
XX
PF 21-DEC-2001; 2001WO-EP015398.
XX
PR 22-DEC-2000; 2000EP-00870318.
PR 04-JAN-2001; 2001EP-00870002.
PR 09-JAN-2001; 2001EP-00870003.
XX
XX
PA (JANC) JANSEN PHARM NV.
XX
PI Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
XX
XX
DR WPI; 2002-667002/71.
XX
XX
DR P-PSDB; ABG93342.
XX
XX
PT New isolated nucleic acid representing a synthetic BAX-gene, useful as
PT medicament for treating, preventing and/or alleviating yeast or fungal
PT infections or proliferative disorders, or for preventing apoptosis in
PT certain diseases.
XX
XX
PS Claim 36; Fig 2; 344pp; English.
XX
XX
CC This invention describes a novel nucleic acid representing a synthetic
CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
CC resistant yeast or fungi, identifying, or obtaining and identifying
CC Candida spp. sequences that are differentially expressed in a pathway
CC eventually leading to programmed cell death or identifying inhibitors or
CC inhibitor sequences of Bax-induced cell death. The products of the
CC invention have cyrostatic, fungicide, immunosuppressive, virucide and
CC vasotropic activity and can be used in vaccines or for gene therapy. The
CC isolated nucleic acids, polypeptides, pharmaceutical compositions,
CC antisense molecules and antibodies are useful as medicaments or in
CC preparing a medicament for treating, preventing and/or alleviating
CC diseases associated with yeast or fungi or proliferative disorders, such
CC as cancer, or for preventing apoptosis in certain diseases. The compounds
CC or polypeptides, or the genetically modified organism are useful for
CC preparing a medicament for modifying the endogenous flora of humans and
CC other mammals. The vaccine is useful for immunising against yeast or
CC fungal infections. Apoptosis-related diseases include autoimmune disease,
CC ischaemia, diseases related with viral infections or neurodegenerations.
CC This sequence represents a polynucleotide associated with the Bax gene
CC described in the disclosure of the invention
XX
XX
SQ Sequence 24076 BP; 7855 A; 4248 C; 4533 G; 7440 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 34 Length: 24076
Score: 70.00 Matches: 10
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 83.33% Indels: 0
DB: 6 Gaps: 0

US-09-858-332g-15 (1-12) x AB076608 (1-24076)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
:::|||||

PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
XX
PS Claim 7; Page 3490; 4525pp; English.

CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and anti-inflammatory
CC activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and
CC antibodies that bind (1) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (1) are used to detect Streptococcus in a
CC biological sample. (1) is used to determine whether a compound binds to
CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (1) may be used to recombinantly produce (1) and may be
CC used in gene therapy. Antibodies to (1) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX

SEQ Sequence 315 BP; 112 A; 55 C; 69 G; 79 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.05 Length: 315
Score: 67.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 79.76% Indels: 0
DB: Gaps: 0

US-09-858-332G-15 (1-12) x ABN67727 (1-315)

OY 1 HisAenHisAenHisAenHisAenHisAenHis 11
DB 7 CATATCATTAACCAAGATCATTAACCAATGAACAT 39

RESULT 50

ABL17932 standard; DNA; 4779 BP.

ABL17932;

26-MAR-2002 (first entry)

Drosophila melanogaster genomic polynucleotide SEQ ID NO 5269.

Drosophila; developmental biology; cell signaling; insecticide;

pharmaceutical; gene; ds.

Drosophila melanogaster.

WO200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001NO-US009231.

23-MAR-2000; 2000US-0191637P.

11-UTL-2000; 2000US-00614150.

(PEKE) PE CORP NY.

Venter JC, Adams M, Li PMD, Myers EW;

WPI; 2001-656860/75.

New isolated nucleic acid detection reagent for detecting 1000 or more

genes from Drosophila and for elucidating cell signaling and cell-cell

PS Claim 1; SEQ ID NO 5269; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL10840-ABL16175) and the encoded proteins (ABBS7737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIFO at ftp.wifo.int/pub/published_pct_sequences
XX

SEQ Sequence 4779 BP; 1534 A; 1148 C; 926 G; 1171 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 16.2 Length: 4779
Score: 67.00 Matches: 9
Percent Similarity: 83.33% Conservative: 1
Best Local Similarity: 75.00% Mismatches: 2
Query Match: 79.76% Indels: 0
DB: Gaps: 0

US-09-858-332G-15 (1-12) x ABL17932 (1-4779)

OY 1 HisAenHisAenHisAenHisAenHisAenHis 12
DB 2475 CACACACACAAACACTCACACACATTAACCAAC 2510

RESULT 51

ABN71527 20/c

Continuation (21 of 22) of ABN71527 from base 2000001 (Streptococcus polynucleotide SEQ

WP Sequence split into 22 fragments LOCUS ABN71527 Accession Abn71527

WP	Fragment Name	Begin	End
WP	ABN71527_00	1	110000
WP	ABN71527_01	100001	210000
WP	ABN71527_02	200001	310000
WP	ABN71527_03	300001	410000
WP	ABN71527_04	400001	510000
WP	ABN71527_05	500001	610000
WP	ABN71527_06	600001	710000
WP	ABN71527_07	700001	810000
WP	ABN71527_08	800001	910000
WP	ABN71527_09	900001	1010000
WP	ABN71527_10	1000001	1110000
WP	ABN71527_11	1100001	1210000
WP	ABN71527_12	1200001	1310000
WP	ABN71527_13	1300001	1410000
WP	ABN71527_14	1400001	1510000
WP	ABN71527_15	1500001	1610000
WP	ABN71527_16	1600001	1710000
WP	ABN71527_17	1700001	1810000
WP	ABN71527_18	1800001	1910000
WP	ABN71527_19	1900001	2010000
WP	ABN71527_20	2000001	210000
WP	ABN71527_21	2100001	2155561

Alignment Scores:

Pred. No.: 384 Length: 110000
Score: 67.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 79.76% Indels: 0
DB: Gaps: 0

US-09-858-332G-15 (1-12) x ABN71527_20 (1-110000)

OY 1 HisAenHisAenHisAenHisAenHisAenHis 11

DB 37697 CATATCATTAACCAAGATCATTAACCAATGAACAT 37665

RESULT 52

AD142533

ID AD142533 standard; DNA; 1338 BP.
 XX
 AC AD142533;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Plant transcription factor polynucleotide #631.
 XX
 KW transgenic; plant; enhanced tolerance to abiotic stress;
 KW glycosylate tolerance; hormone sensitivity; disease resistance;
 KW sugar sensing; flowering; flower structure; stem bifurcation;
 KW branching pattern; apical dominance; trichome; stem morphology;
 KW root growth; root hair; seed development; cell proliferation;
 KW cell differentiation; premature senescence; necrosis; plant size;
 KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
 KW plant anthocyanin; light response; shade avoidance; bioinformatic;
 KW transcription factor; gene; ds.
 XX
 OS Glycine max.
 XX
 PN US2004019927-A1.
 XX
 PD 29-JAN-2004.
 XX
 PF 25-FEB-2003; 2003US-00374780.
 XX
 PR 18-APR-2001; 2001US-00837944.
 XX
 PA (SHER/) SHERMAN B K.
 PA (RIEC/) RIECHMANN J L.
 PA (JIANG/) JIANG C.
 PA (HEAR/) HEARD J E.
 PA (HAAK/) HAAKE V.
 PA (CREE/) CREELMAN R A.
 PA (RATC/) RATCLIFFE O.
 PA (ADAM/) ADAM L J.
 PA (REUB/) REUBER T L.
 PA (KEDD/) KEDDIE J.
 PA (BROU/) BROUN P E.
 PA (PIIG/) PIIGRIM M L.
 PA (DUBE/) DUBBEL A N.
 PA (PINE/) PINEDA O.
 PA (YUGG/) YU G.
 XX
 PI Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
 PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;
 PI Pigrim ML, Dubbel AN, Pineda O, Yu G;
 XX
 DR WPI; 2004-132245/13.
 XX
 PT New transgenic plant comprising a recombinant polynucleotide of any one
 PT of more than 500 nucleotide sequences, useful in bioinformatic search
 PT methods.
 XX
 PS Claim 1; SEQ ID NO 996; 435bp; English.
 XX
 CC The invention describes a transgenic plant comprising a recombinant
 CC polynucleotide of any one of more than 500 nucleotide sequences fully
 CC defined in the specification or its complement. The method of the
 CC invention can be used to produce a plant having altered traits such as:
 CC enhanced tolerance to abiotic stress; glycosylate tolerance; hormone
 CC sensitivity; disease resistance; sugar sensing; early or late flowering;
 CC altered flower structure; change in stem bifurcations; altered branching
 CC pattern; reduced apical dominance; reduced trichome density; lack of
 CC trichomes; reduced ectopic trichome development; altered trichome
 CC development; increase in trichome number; altered stem morphology;
 CC increased root growth; increased root hairs; altered seed development;
 CC altered cell proliferation or cell differentiation; rapid development;
 CC premature senescence; increased necrosis; increase in seedling or plant
 CC size; decreased plant size; leaf morphology; seed morphology; seed
 CC biochemistry; increase in root anthocyanins; increase in plant
 CC anthocyanins; or alteration in light response or shade avoidance. The
 CC transgenic plant, polynucleotides and polypeptides are useful in

CC bioinformatic search methods. This sequence represents a plant
 CC transcription factor, and an orthologue of Arabidopsis thaliana
 CC transcription factor isolated in the invention, that can be used in the
 CC creation of a transgenic plant with altered traits.
 XX
 SQ Sequence 1338 BP; 449 A; 220 C; 370 G; 299 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 6.05 Length: 1338
 Score: 66.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 2
 Best Local Similarity: 81.82% Mismatches: 0
 Query Match: 78.57% Indels: 0
 DB: 12 Gaps: 0
 US-09-858-332g-15 (1-12) x AD142533 (1-1338)
 QY 1 H1SA8NH1SA8NH1SA8NH1SA8NH1SA8NH1S 11
 Db 54 TACAACCAACCAACCACTACAACCAACCAAC 86
 RESULT 53
 ID ADO02862 standard; cDNA; 1338 BP.
 XX
 AC ADO02862;
 XX
 DT 01-UTL-2004 (first entry)
 XX
 DE Soybean orthologue of Thalecress transcription factor, cDNA #139.
 XX
 KW Soybean; transcription factor; ss; gene; plant; transgenic;
 KW abiotic stress; cold tolerance; heat tolerance; drought; osmotic stress;
 KW phosphate limitation; potassium limitation; sugar sensing; seed germination;
 KW hormone sensitivity; disease resistance; apical dominance;
 KW flowering; inflorescence architectural change;
 KW meristem cell differentiation; phytoalexin; apical dominance;
 KW trichome development; seed development; premature senescence;
 KW delayed senescence; lethality; necrosis; plant size; leaf morphology;
 KW seed morphology; secondary metabolism; light response; shade avoidance.
 XX
 OS Glycine max.
 XX
 PN US2004045049-A1.
 XX
 PD 04-MAR-2004.
 XX
 PF 10-APR-2003; 2003US-00412699.
 XX
 PR 13-SEP-1999; 99US-00394519.
 PR 21-JAN-2000; 2000US-00489376.
 PR 17-FEB-2000; 2000US-00506720.
 PR 22-MAR-2000; 2000US-00532591.
 PR 22-MAR-2000; 2000US-00533029.
 PR 22-MAR-2000; 2000US-00533030.
 PR 22-MAR-2000; 2000US-00533392.
 PR 22-MAR-2000; 2000US-00533648.
 PR 06-APR-2000; 2000WO-US009948.
 PR 16-NOV-2000; 2000US-00713994.
 PR 27-MAR-2001; 2001US-00819142.
 PR 17-APR-2001; 2001US-00837444.
 PR 30-JAN-2002; 2002US-00958131.
 PR 14-JUN-2002; 2002US-00171468.
 PR 09-AUG-2002; 2002US-00225066.
 PR 09-AUG-2002; 2002US-00225067.
 PR 09-AUG-2002; 2002US-00225068.
 PR 17-DEC-2002; 2002US-0434166P.
 PR 25-FEB-2003; 2003US-00374780.
 XX
 PA (ZHAN/) ZHANG J.
 PA (FROM/) FROMM M E.
 PA (HEAR/) HEARD J E.
 PA (RIEC/) RIECHMANN J L.

Alignment Scores:

Pred. No.: 26.7 Length: 5826
 Score: 66.00 Matches: 9
 Percent Similarity: 91.67% Conservative: 2
 Best Local Similarity: 75.00% Mismatches: 1
 Query Match: 78.57% Indels: 0
 DB: 6 Gaps: 0

US-09-858-332g-15 (1-12) x ABLJ3562 (1-5826)

QY 1 HisAgnHisAgnHisAgnHisAgnHisAgn 12

DB 5525 TATACCATTAACCATTAACCATTAACCATTAAT 5490

RESULT 55

ID ABLJ3631/C standard, DNA; 6236 BP.

AC ABLJ3631;

DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 1604.

XX Human; immune system disease; cytosine methylation; antiasthmatic;

XX antiarteriosclerotic; anti-naemic; cytosine; noctropic;

XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;

XX antineumatic; antirheumatic; antidiabetic; antiparotitic;

XX antineumatic; cancer; eye disease; arteriosclerosis; anaemia;

XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;

XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;

XX ds.

XX Homo sapiens.

XX WO200200928-A2.

XX PD 03-JAN-2002.

XX PF 02-JUL-2001; 2001WO-EP007537.

XX PR 30-JUN-2000; 2000DE-01032529.

XX PT 01-SEP-2000; 2000DE-01043826.

XX PA (EPIC-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX DR WPI; 2002-130909/17.

XX PT Nucleic acid comprising fragment of chemically modified gene, useful for

XX diagnosis and treatment of diseases associated with abnormal cytosine

XX methylation.

XX Claim 1; SEQ ID NO 1604; 32pp + Sequence Listing; German.

XX CC The present invention provides a number of human immune system associated

XX genes which are modified by the methylation of cytosine. The sequences

XX can be used in the diagnosis and treatment of immune system disorders,

XX including eye diseases such as retinopathy, neovascular glaucoma and

XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid

XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,

XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel

XX diseases. The present sequence is a gene of the invention

XX SQ Sequence 6236 BP; 1352 A; 206 C; 1786 G; 2892 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 28.6 Length: 6236

XX Score: 66.00 Matches: 9

XX Percent Similarity: 91.67% Conservative: 2

XX Best Local Similarity: 75.00% Mismatches: 1

Query Match: 78.57% Indels: 0
 DB: 6 Gaps: 0

US-09-858-332g-15 (1-12) x ABLJ3631 (1-6236)

QY 1 HisAgnHisAgnHisAgnHisAgnHisAgn 12

DB 2029 TATACCATTAACCATTAACCATTAACCATTAAT 1994

RESULT 56

ID ABLJ3631/C standard, DNA; 6236 BP.

AC ABLJ3631;

DT 15-JUL-2002 (first entry)

DE Human chemically modified disease associated gene SEQ ID NO 220.

XX Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;

XX heart disease; epilepsy; histone deacetylation; muscular dystrophy;

XX dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;

XX antidiabetic; cytosine; anticonvulsant; ds.

XX Homo sapiens.

XX Synthetic.

XX WO200200927-A2.

XX PD 03-JAN-2002.

XX PF 02-JUL-2001; 2001WO-EP007536.

XX PR 30-JUN-2000; 2000DE-01032529.

XX PT 01-SEP-2000; 2000DE-01043826.

XX PA (EPIC-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX DR WPI; 2002-130908/17.

XX PT Novel nucleic acid useful for diagnosis and therapy of diseases

XX associated with development genes such as diabetes, comprises a sequence

XX of a segment of chemically pretreated DNA of genes associated with

XX development.

XX Claim 1; SEQ ID NO 220; 27pp; English.

XX CC The invention relates to a nucleic acid (I) comprising a sequence at

XX least 18 bases in length of a segment of chemically pretreated DNA (II)

XX of genes associated with development selected from 87 genes listed in the

XX sequences (ABN79984-ABN80333) or their complements. The invention is

XX useful for the diagnosis or therapy of diseases associated with

XX development genes, in particular disease related to homeobox containing

XX genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes

XX associated with congenital heart disease, epilepsy, diseases related to

XX histone deacetylation, Curran's syndrome, diseases related with the

XX development of the brain and limb girdle muscular dystrophy and dwarfism.

XX Oligomers specific to each of the genes are useful for detecting the

XX methylation state of all CpG dinucleotides within the 350 sequences or

XX (II) and their complementary sequences, as primer oligonucleotides for

XX the amplification of the 350 sequences, (II) and/or their complements and

XX as oligomer probes for detecting the cytosine methylation state and/or

XX single nucleotide polymorphisms (SNPs). Note: The sequence data for this

XX patent did not form part of the printed specification but is based on

XX sequence information supplied to Derwent by the European Patent Office

XX SQ Sequence 6236 BP; 1352 A; 206 C; 1786 G; 2892 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 28.6 Length: 6236

Score: 66.00 Matches: 9
 Percent Similarity: 91.67% Conservative: 2
 Best Local Similarity: 75.00% Mismatches: 1
 Query Match: 78.57% Indels: 0
 DB: 6 Gaps: 0

US-09-858-332G-15 (1-12) x ABN80203 (1-6236)

OY 1 HisAnH1sAAnH1sAAnH1sAAnH1sAAn 12
 Db 2029 TATACCATTAACCATTAACCATTAACCATTAAT 1994

RESULT 57
 ID ABL06142 standard; cDNA; 28360 BP.
 AC ABL06142;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 12908.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN MO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW,
 XX
 DR WPI; 2001-656660/75.
 XX
 DR P-PSDB; ABB62039.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Claim 1; SEQ ID NO 12908; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB7072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 28360 BP; 7506 A; 6320 C; 6583 G; 7951 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 132 Length: 28360
 Score: 66.00 Matches: 10
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 2
 Query Match: 78.57% Indels: 0
 DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x ABL06142 (1-28360)

OY 1 HisAnH1sAAnH1sAAnH1sAAnH1sAAn 12

Db 22869 CACAAATCAATCAATCAATCAATCAATCAATCAAT 22834

RESULT 58

ID ABS69323 standard; DNA; 334 BP.

AC ABS69323;
 XX
 DT 21-NOV-2002 (first entry)
 XX
 DE Novel murine polynucleotide isolated using gene trap technology #386.
 XX
 KW Mouse; gene trapped sequence; GTS; functional genomic analysis;
 KW phage display system; gene chip; temporal gene expression;
 KW tissue specific gene expression; antisense inhibition; gene targeting;
 KW development disorder; cell differentiation disorder; aging; cancer;
 KW autoimmune disease; lupus; inflammatory disorder; skin disorder;
 KW degenerative disorder; ds.
 XX
 OS Mus musculus.
 XX
 PN US2002102543-A1.
 XX
 PD 01-AUG-2002.
 XX
 PF 30-NOV-2000; 2000US-00728445.
 XX
 PR 01-DEC-1999; 99US-0168358P.
 XX
 PA (FRIE/) FRIEDRICH G.
 PA (ZAMB/) ZAMBROWICZ B.
 PA (SAND/) SANDS A T.
 XX
 PI Friedrich G, Zambrowicz B, Sands AT;
 XX
 DR WPI; 2002-690598/74.
 XX
 PT Novel murine polynucleotides that individually identify novel genes into
 PT which a retroviral gene trap vector has integrated, useful in genomic
 PT analysis and in discovery, development of therapeutic and diagnostic
 PT agents.
 XX
 PS Claim 1; Page 138; 296pp; English.
 XX
 CC The invention describes an isolated murine polynucleotide (I) comprising
 CC a contiguous stretch of at least 60 nucleotides of one of 265-677
 CC nucleotide 891 OMNIBANK gene trapped sequences (GTSs) (S), given in the
 CC specification. The novel genes and cells are useful in functional genomic
 CC analysis and in the discovery and development of new therapeutic and
 CC diagnostic agents and methods. (I) is useful for identifying the coding
 CC regions of the murine genome, to isolate cDNAs, genomic clones, or full-
 CC length genes/polynucleotides or homologues, heterologues, paralogues, or
 CC orthologues that are capable of hybridizing to one or more of the GTSs
 CC under stringent conditions. (I) can be incorporated into a phage display
 CC system that can be used to screen for proteins, or other ligands, that
 CC are capable of binding an amino acid sequence encoded by an
 CC oligonucleotide or polynucleotide sequence in at least one of the TS
 CC sequences. (I) is useful in addressable arrays, such as gene chips, to
 CC identify and characterize temporal and tissue specific gene expression,
 CC to identify the gene of interest from many sources and for genetic
 CC manipulations such as antisense inhibition and gene targeting. Decreasing
 CC the level of expression of (I) and/or down regulating the activity of
 CC peptides or proteins encoded by (I) is useful for treating development
 CC and cell differentiation disorders, aging, cancer, autoimmune disease,
 CC lupus, inflammatory disorders, skin disorders and degenerative disorders.
 CC This sequence represents a murine cDNA isolated using gene trap
 CC technology
 XX
 SQ Sequence 334 BP; 109 A; 86 C; 96 G; 43 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.01 Length: 334

Alignment Scores:

Pred. No.:	4.24	length:	700
Score:	65.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	3
Best Local Similarity:	75.00%	Mismatches:	0
Query Match:	77.38%	Indels:	0
DB:	4	Gaps:	0

US-09-858-332G-15 (1-12) x AAF23528 (1-700)

[illegible]

```

Oy      1   Hisasnihsasnhisasnhisasnhisasn 12
          |||||
Db      605 CACAGCACCAACTGCACCGGACGCCACACACACAC 646
          |||||

RESULT: 62
ABL27328
ID      ABL27328 standard; DNA; 1958 BP.

```

Alignment Scores:	
pred. No.:	5.62
Score:	65.00
Percent Similarity:	83.33%
Best Local Similarity:	83.33%
Query Match:	77.38%
DB:	4
US-09-858-332G-15 (1-12) × AB277329 (1-927)	
	length: 927
	Matches: 10
	Conservative: 6
	Mismatches: 2
	Indels: 0
	Gaps: 0

QY	1	His	asn	his	asn	his	asn	his	asn	his	asn	asn	12
Db	636	C	A	C	A	C	A	C	A	C	A	C	671
RESULT 63													
ID	AAV20700	standard; DNA; 5163 BP.											
XX	AAV20700;												
AC	AAV20700;												

DT		XX	17-AUG-1998	(first entry)	
XX		XX	Cryptosporidium parvum GP900 antigen open reading frame.		
DE		XX	Open reading frame; ORF; antigen; GP900; cryptosporidium; infection;		
XX		KW	antibody; propylaxis; treatment; inhibition; retardation; detection;		
KV		KW	diagnosis; human; ds.		
XX		OS	Cryptosporidium parvum.		
FT		XX			
FH		XX			
Key		XX	Location/Qualifiers		
CDS		FT	1..5163		
FT		/+tag=	a		
misc_feature		/product=	"GP900 antigen"		
FT		524..1270			
FT		/+tag=	b		
FT		/note=	"region containing NINC mutations"		
XX					
PN		XX	M09806430-A1.		
PD		XX	19-FEB-1998.		
PF		XX	11-AUG-1997; 97WO-US014104.		
PR		XX	14-AUG-1996; 96US-00700651.		
PA		XX	(REGC) UNIV CALIFORNIA.		
P1		XX	Petersen C, Leech J, Nelson RC, Gut J;		
PI		XX	WPI; 1998-159290/14.		
DR		XX	P-PADB; AAM48299.		
PT		XX	Anti-Cryptosporidium antibody - used to develop products for detection,		
PP		XX	diagnosis, propylaxis or treatment of Cryptosporidium infections.		
PS		XX	Claim 32; Page 60-62; 89pp; English.		
CC		XX	The sequence is that encoding the GP900 antigen which may be used in the		
CC		XX	production of anti-Cryptosporidium antibodies. These can be used for the		
CC		XX	propylaxis, treatment, inhibition or retardation of a Cryptosporidium		
CC		XX	infection in humans or in animals such as calves. They can also be used		
CC		XX	for the detection and diagnosis of related infections		
SQ		XX	Sequence 5163 BP; 1875 A; 1137 C; 873 G; 1278 T; 0 U; 0 Other;		
Alignment Scores:					
Pred. NO.:	31.8	Length:	5163		
Score:	65.00	Matches:	11		
Percent Similarity:	66.67%	Conservative:	1		
Best local Similarity:	61.11%	Mismatches:	0		
Query Match:	77.38%	Indels:	6		
	2	Gaps:	1		
US-09-858-332G-15 (1-12) x AAV20700 (1-5163)					
Oy	1 HisAsnHisAsnHis-----AsnHisAsnHisAsnHisAsn 12				
Dd	:::				
	856 CACACACTCACACAACCACAACTACACACACACACACACACCAACCAAC 909				
RESULT 64					
ID	AAA61849 standard; DNA; 5163 BP.				
XX	AAA61849;				
AC					
DT	28-OCT-2000 (first entry)				
ORF	encoding a portion of Cryptosporidium parvum NINC isolate GP900.				
NINC	isolate; glycoprotein; antibody; cryptosporidiosis;				
competitive	inhibition; attachment; invasion; ligand binding; sporozoite;				

XX		merozoite; diarrhoea; protozoacide; open reading frame; ORF; db.
XX	Cryptosporidium parvum.	
OS		
FH	Key	Location/Qualifiers
FT	CDS	1669..7182
FT		/tag= a
FT		/partial
FT		/product= "Cryptosporidium parvum NINC isolate GP900"
FT		/note= "No start or stop codons given in the
FT		specification"
XX		
XX	US6071518-A.	
PN		
PD	06-JUN-2000.	
XX		
FF	12-SEP-1997;	97US-00928361.
XX		
PR	29-MAY-1992;	92US-00891301.
PR	01-JUN-1993;	93US-00071880.
PR	03-APR-1995;	95US-00415751.
PR	14-AUG-1996;	96US-00700651.
PR	13-SEP-1996;	96US-0026062P.
XX		
PA	(REGC) UNIV CALIFORNIA.	
PI	Petersen C;	
PS	WPI, 2000-422065/36.	
DR	P-PBDS; AAB11727.	
XX		
PT	New GP900 protein fragments and fusion proteins of Cryptosporidium	
PT	parvum, useful for detecting the presence of the parasite, and diagnosing	
PT	or treating Cryptosporidium infections by competitive inhibition of the	
PT	function of GP900.	
XX		
XX	Claim 16; Col 47-52; 59pp; English.	
CC	The invention relates to the GP900 glycoprotein of the protozoan	
CC	Cryptosporidium parvum, DNA encoding it, GP900 fragments, and fusion	
CC	proteins comprising GP900 fragments. The invention also relates to the	
CC	administration of GP900 or fragments thereof to a host to elicit anti-	
CC	GP900 antibody production, and to a method of cryptosporidiosis treatment	
CC	or prophylaxis comprising administration of anti-GP900 antibodies to an	
CC	individual. Cryptosporidium parvum GP900 and GP900 fragments are able to	
CC	competitively inhibit sporozoite or merozoite attachment or invasion, and	
CC	are also useful for the generation of anti-GP900 antibodies. The	
CC	antibodies also inhibit sporozoite or merozoite attachment/invasion and	
CC	additionally inhibit the binding of GP900 ligands to GP900. GP900	
CC	proteins, fragments and antibodies may therefore be used to treat or	
CC	prevent cryptosporidiosis. Infection with Cryptosporidium is a common	
CC	cause of diarrhoea in humans and causes life-threatening diarrhoea in	
CC	immunocompromised persons. Cryptosporidiosis can be contracted from	
CC	contaminated municipal water supplies (e.g., public swimming pools). It	
CC	is also a cause of disease in animals, resulting in financial losses in	
CC	agriculture. GP900 fragments, fusion proteins and antibodies may also be	
CC	used for the diagnosis of Cryptosporidium parvum infections, and for the	
CC	detection of the parasite in the environment. The present sequence	
CC	represents the open reading frame (ORF) encoding a portion of the GP900	
CC	protein of the NINC isolate of Cryptosporidium parvum	
XX		
SO	Sequence 5163 BP; 1873 A; 1138 C; 875 G; 1277 T; 0 U; 0 Other;	
	Alignment Scores:	
	Pred. No.:	31.8
	Score:	65.00
	Percent Similarity:	66.67%
	Best Local Similarity:	61.11%
	Query Match:	77.38%
	DB:	3
	Gaps:	1
	Length:	5163
	Matches:	11
	Conservative:	1
	Mismatches:	0
	Indels:	6
	Gaps:	1
	US-09-858-3326-15 (1-12) x AAA61849 (1-5163)	

OY 1 HisaHnHsAenHs-----AenHsAenHsAenHsAen 12
DB 855 CACAACCTACCAACCAACCACTACCAACCAACCAACCAACCAACCAAC 908
RESULT 65
ID ABR04778 standard; DNA; 5163 BP.
XX ABR04778;
AC ABR04778;
XX
XX 27-SEP-2002 (first entry)
DT
XX
DE C parvum GP900 gene fragment SEQ ID NO: 4.
XX
XX
XX CRYPTOSPORIDIUM detection; GP900; P68; cryptopain; cryptosporidiosis;
XX gene; ds.
XX
XX CRYPTOSPORIDIUM parvum.
XX
XX MO200194631-A1.
XX
XX 13-DEC-2001.
XX
XX 14-MAY-2001; 2001WO-US015624.
XX
XX 06-JUN-2000; 2000US-00588995.
XX
XX (REGC) UNITV CALIFORNIA.
XX
XX Petersen C, Barnes DA, Nelson RG, Gut J;
XX WPI; 2002-566447/60.
XX
XX
XX Detecting Cryptosporidium in biological and environmental samples and
PT diagnosis of cryptosporidiosis involves, contacting the sample with
PT CRYPTOSPORIDIUM GP900, P68 or cryptopain antigen, antibody, DNA or RNA.
XX
XX
XX Disclosure; Page 103-104; 157pp; English.
XX
XX The present invention relates to a method of detecting Cryptosporidium in
CC biological and environmental samples, and of diagnosing
CC cryptosporidiosis. This involves obtaining a sample and contacting it
CC with CRYPTOSPORIDIUM GP900, P68 or cryptopain antigen, antibody, DNA or
CC RNA, or its variant, mutant or fragment. The method is also useful for
CC detecting and identifying individual Cryptosporidium isolates based on
CC the genetic characteristics, and for diagnosis of prior or concurrent
CC cryptosporidium infection. The present sequence is a C. parvum coding
CC sequence used in the exemplification of the invention
XX
XX
SQ Sequence 5163 BP; 1873 A; 1138 C; 875 G; 1277 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
Pred. No.: 31.8 Length: 5163
Score: 65.00 Matches: 11
Percent Similarity: 66.67% Conservative: 1
Best Local Similarity: 61.1% Mismatches: 0
Query Match: 77.38% Indels: 6
DB: Gaps: 1
XX
XX US-09-858-332G-15 (1-12) x ABR04778 (1-5163)
OY 1 HisaHnHsAenHs-----AenHsAenHsAenHsAen 12
DB 855 CACAACCTACCAACCAACCACTACCAACCAACCAACCAACCAACCAAC 908
RESULT 66
ID AAV20701
XX AAV20701 standard; DNA; 5318 BP.
XX
XX AAV20701;
XX
XX 17-AUG-1998 (first entry)
XX

DE CRYPTOSPORIDIUM parvum GP900 antigen open reading frame and 3' region.
XX
XX Open reading frame; ORF; antigen; GP900; cryptosporidium; infection;
XX antibody; prophylaxis; treatment; inhibition; retardation; detection;
XX diagnosis; human; 3' region; ds.
XX
XX CRYPTOSPORIDIUM parvum.
XX
XX
XX Key Location/Qualifiers
FH 1..5166
FT CDS /tag= a
FT /product= "GP900 antigen"
FT misc_feature 524..1270
FT /tag= b
XX /note= "region containing NINC mutations"
XX
XX MO9806430-A1.
XX
XX 19-FEB-1998.
XX
XX 11-AUG-1997; 97WO-US014104.
XX
XX 14-AUG-1996; 96US-00700651.
XX
XX (REGC) UNITV CALIFORNIA.
XX
XX Petersen C, Leech J, Nelson RC, Gut J;
XX WPI; 1998-159290/14.
XX P-PSDB; AAW48299.
XX
XX
XX Anti-Cryptosporidium antibody - used to develop products for detection,
PT diagnosis, prophylaxis or treatment of Cryptosporidium infections.
XX
XX
XX Claim 32; Page 62-63; 89pp; English.
XX
XX The sequence is that encoding the GP900 antigen which may be used in the
CC production of anti-Cryptosporidium antibodies. These can be used for the
CC prophylaxis, treatment, inhibition or retardation of a Cryptosporidium
CC infection in humans or in animals such as calves. They can also be used
CC for the detection and diagnosis of related infections
XX
XX
SQ Sequence 5318 BP; 1944 A; 1153 C; 888 G; 1333 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
Pred. No.: 32.7 Length: 5318
Score: 65.00 Matches: 11
Percent Similarity: 66.67% Conservative: 1
Best Local Similarity: 61.1% Mismatches: 0
Query Match: 77.38% Indels: 6
DB: Gaps: 1
XX
XX US-09-858-332G-15 (1-12) x AAV20701 (1-5318)
OY 1 HisaHnHsAenHs-----AenHsAenHsAenHsAen 12
DB 856 CACAACCTACCAACCAACCACTACCAACCAACCAACCAACCAACCAAC 909
RESULT 67
ID AAA61848
XX AAA61848 standard; DNA; 5318 BP.
XX
XX AAA61848;
XX
XX 28-OCT-2000 (first entry)
DT
XX
XX DNA encoding a portion of Cryptosporidium parvum NINC isolate GP900.
DE
XX
XX GP900; NINC isolate; glycoprotein; antibody; cryptosporidiosis;
XX competitive inhibition; attachment; invasion; ligand binding; sporozoite;
XX merozoite; diarrhoea; protozoacide; ds.
XX
XX CRYPTOSPORIDIUM parvum.
OS

KM Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX Drosophila melanogaster.
XX WO200171042-A2.
PD 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009221.
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI: 2001-6556860/75.
DR P-PSDB; ABB6951.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX Claim 1; SEQ ID NO 36644; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 13033 BP; 3717 A; 3149 C; 2905 G; 3262 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 80.9 Length: 13033
Score: 65.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 77.38% Indels: 0
Gaps: 0
US-09-858-332g-15 (1-12) x ABL14054 (1-13033)
QY 1 HisAenHisAenHisAenHisAenHisAenHis 11
DB 8114 CACNACCACNACCACNACCACNACCACNACCAC 8082
RESULT 70
AB067196_5
Continuation (6 of 7) of AB067196 from base 500001 (Listeria innocua contig DNA sequence
WP Sequence split into 7 fragments LOCUS AB067196 Accession Abq67196
WP Fragment Name Begin End
WP AB067196_0 1 110000
WP AB067196_1 100001 210000
WP AB067196_2 200001 310000
WP AB067196_3 300001 410000
WP AB067196_4 400001 510000
WP AB067196_5 500001 610000
WP AB067196_6 600001 684707
Alignment Scores:
Pred. No.: 695 Length: 110000
Score: 65.00 Matches: 9
Percent Similarity: 83.33% Conservative: 1
Best Local Similarity: 75.00% Mismatches: 2
Query Match: 77.38% Indels: 0
Gaps: 0

DB: 6 Gaps: 0
US-09-858-332g-15 (1-12) x AB067196_5 (1-110000)
QY 1 HisAenHisAenHisAenHisAenHisAenHis 12
DB 42149 CATNATCAGCATCATGTCACATGTCACATNATC 42184
RESULT 71
AB069245_27/c
Continuation (28 of 31) of AB069245 from base 2700001 (Listeria innocua DNA sequence #6
WP Sequence split into 31 fragments LOCUS AB069245 Accession Abq69245
WP Fragment Name Begin End
WP AB069245_00 1 110000
WP AB069245_01 100001 210000
WP AB069245_02 200001 310000
WP AB069245_03 300001 410000
WP AB069245_04 400001 510000
WP AB069245_05 500001 610000
WP AB069245_06 600001 710000
WP AB069245_07 700001 810000
WP AB069245_08 800001 910000
WP AB069245_09 900001 1010000
WP AB069245_10 1000001 1110000
WP AB069245_11 1100001 1210000
WP AB069245_12 1200001 1310000
WP AB069245_13 1300001 1410000
WP AB069245_14 1400001 1510000
WP AB069245_15 1500001 1610000
WP AB069245_16 1600001 1710000
WP AB069245_17 1700001 1810000
WP AB069245_18 1800001 1910000
WP AB069245_19 1900001 2010000
WP AB069245_20 2000001 2110000
WP AB069245_21 2100001 2210000
WP AB069245_22 2200001 2310000
WP AB069245_23 2300001 2410000
WP AB069245_24 2400001 2510000
WP AB069245_25 2500001 2610000
WP AB069245_26 2600001 2710000
WP AB069245_27 2700001 2810000
WP AB069245_28 2800001 2910000
WP AB069245_29 2900001 3010000
WP AB069245_30 3000001 3011208
Alignment Scores:
Pred. No.: 695 Length: 110000
Score: 65.00 Matches: 9
Percent Similarity: 83.33% Conservative: 1
Best Local Similarity: 75.00% Mismatches: 2
Query Match: 77.38% Indels: 0
Gaps: 0
US-09-858-332g-15 (1-12) x AB069245_27 (1-110000)
QY 1 HisAenHisAenHisAenHisAenHisAenHis 12
DB 38663 CATNATCAGCATCATGTCACATGTCACATNATC 38628
RESULT 72
ABA03041_26/c
Continuation (27 of 30) of ABA03041 from base 2600001 (Listeria monocytogenes EGD-e geno
WP Sequence split into 30 fragments LOCUS ABA03041 Accession ABA03041
WP Fragment Name Begin End
WP ABA03041_00 1 110000
WP ABA03041_01 100001 210000
WP ABA03041_02 200001 310000
WP ABA03041_03 300001 410000
WP ABA03041_04 400001 510000
WP ABA03041_05 500001 610000
WP ABA03041_06 600001 710000
WP ABA03041_07 700001 810000
WP ABA03041_08 800001 910000

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 25; SEQ ID NO 11701; 487bp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 148 BP; 60 A; 56 C; 14 G; 18 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,19 Length: 148
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 4 Gaps: 0
XX
US-09-858-332g-15 (1-12) x AA121768 (1-148)
XX
QY 1 HisaSnH1sAsnH1sAsnH1sAsnH1sAsnH1sAsn 12
DB 11 CACAACCAACCAACCAACCAACCAACCAACCAACCAAC 46
XX
RESULT 75
ABA66842
ID ABA66842 standard; DNA; 148 BP.
XX
AC ABA66842;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #15147.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human foetal liver.
XX

PS Claim 4; SEQ ID NO 15147; 639bp + Sequence Listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human foetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 148 BP; 60 A; 56 C; 14 G; 18 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,19 Length: 148
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 4 Gaps: 0
XX
US-09-858-332g-15 (1-12) x ABA66842 (1-148)
XX
QY 1 HisaSnH1sAsnH1sAsnH1sAsnH1sAsnH1sAsn 12
DB 11 CACAACCAACCAACCAACCAACCAACCAACCAACCAAC 46
XX
RESULT 76
AA147050
ID AA147050 standard; DNA; 148 BP.
XX
AC AA147050;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #15736 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 15736; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENPs).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX

SQ Sequence 148 BP; 60 A; 56 C; 14 G; 18 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.19 Length: 148
 Score: 64.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 4
 Best Local Similarity: 66.67% Mismatches: 0
 Query Match: 76.19% Indels: 0
 DB: 4 Gaps: 0
 US-09-858-332g-15 (1-12) x AAI47050 (1-148)
 QY 1 H1sAsnH1sAsnH1sAsnH1sAsnH1sAsnH1sAsn 12
 Db 11 CACAAACCAACCAACCAACCAACCAACCAACCAAC 46
 RESULT 77
 ID ABA48927 standard; DNA; 148 BP.
 AC ABA48927;
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Human breast cell single exon nucleic acid probe #7622.
 XX
 KW Human; microarray; single exon probe; gene expression; breast; disease;
 cancer; ss.
 OS Homo sapiens.
 XX
 PN WO200157271-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000662.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 XX
 PR 26-MAY-2000; 2000US-0207456P.
 XX
 PR 30-JUN-2000; 2000US-00608408.
 XX
 PR 03-AUG-2000; 2000US-00632366.
 XX
 PR 21-SEP-2000; 2000US-0234687P.
 XX
 PR 27-SEP-2000; 2000US-0236359P.
 XX
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-496933/54.
 XX
 PT New spatially-addressable set of single exon nucleic acid probes, useful
 PT for measuring gene expression in sample derived from human breast,
 PT comprises number of single exon nucleic acid probes.
 XX
 PS Claim 4; SEQ ID NO 7622; 327bp + Sequence Listing; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and Bt 474 cells. The method involves contacting the
 CC probes with a collection of detectably labelled nucleic acids derived
 CC from mRNA of human breast, and then measuring the label bound to each
 CC probe of the microarray. The probes are useful for verifying the
 CC expression of regions of genomic DNA predicted to encode proteins. They
 CC are useful for gene discovery, and for determining predisposition and/or
 CC prognosing breast disease. Gene expression analysis is useful for
 CC assessing the toxicity of chemical agents on cells. The microarray of
 CC this invention presents a far greater diversity of probes for measuring
 CC gene expression, with far less bias than expressed sequence tag
 CC microarrays. The method is suitable for rapid production of functional
 CC information from genomic sequence. The present sequence is a single exon
 CC nucleic acid probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 148 BP; 60 A; 56 C; 14 G; 18 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.19 Length: 148
 Score: 64.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 4
 Best Local Similarity: 66.67% Mismatches: 0
 Query Match: 76.19% Indels: 0
 DB: 4 Gaps: 0
 US-09-858-332g-15 (1-12) x ABA48927 (1-148)
 QY 1 H1sAsnH1sAsnH1sAsnH1sAsnH1sAsnH1sAsn 12
 Db 11 CACAAACCAACCAACCAACCAACCAACCAACCAAC 46
 RESULT 78
 ID ABA33909 standard; DNA; 148 BP.
 AC ABA33909;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Probe #12375 for gene expression analysis in human heart cell sample.
 XX
 KW Human; gene expression; heart; microarray; vascular system; probe;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 congenital heart disease; ss.
 OS Homo sapiens.
 XX
 PN WO200157274-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000666.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 XX
 PR 26-MAY-2000; 2000US-0207456P.
 XX
 PR 30-JUN-2000; 2000US-00608408.
 XX
 PR 03-AUG-2000; 2000US-00632366.
 XX
 PR 21-SEP-2000; 2000US-0234687P.
 XX
 PR 27-SEP-2000; 2000US-0236359P.
 XX
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488899/53.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts.
 XX
 PS Claim 4; SEQ ID NO 12375; 530bp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart. The
 CC present sequence is one such probe. The probes may be used for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from the human heart via microarrays. By measuring gene expression, the
 CC probes are useful for predicting, diagnosing, grading, staging,
 CC monitoring and prognosing diseases of the human heart and vascular system
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
 CC congenital heart disease. Note: The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 148 BP; 60 A; 56 C; 14 G; 18 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.19 Length: 148
 Score: 64.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 4
 Best Local Similarity: 66.67% Mismatches: 0
 Query Match: 76.19% Indels: 0
 DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x ABA33909 (1-148)

QY 1 HisAenHISAenHISAenHISAenHISAen 12
 Db 11 CACAACCAACACCAACCAACCAACCAACCAAC 46

RESULT 79

AAK40999
 ID AAK40999 standard; DNA; 148 BP.

AC AAK40999;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 15556.

KM Human; bone marrow expressed exon; gene expression analysis; probe;
 KM microarray; cancer; leukaemia; lymphoma; myeloma; ss.

OS Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000668.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-0063236P.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human bone marrow.
 PS Example 4; SEQ ID NO 15556; 658bp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention

SQ Sequence 148 BP; 60 A; 56 C; 14 G; 18 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.19 Length: 148
 Score: 64.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 4
 Best Local Similarity: 66.67% Mismatches: 0
 Query Match: 76.19% Indels: 0
 DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x AAK40999 (1-148)

QY 1 HisAenHISAenHISAenHISAenHISAen 12
 Db 11 CACAACCAACCAACCAACCAACCAACCAACCAAC 46

RESULT 80

AAK15273
 ID AAK15273 standard; DNA; 148 BP.

AC AAK15273;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe SEQ ID NO: 15264.

KM Human; brain expressed exon; gene expression analysis; probe; microarray;
 KM Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
 KM ss.

OS Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000667.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-0063236P.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.
 PS Example 4; SEQ ID NO 15264; 650pp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention

SQ Sequence 148 BP; 60 A; 56 C; 14 G; 18 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.19 Length: 148
 Score: 64.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 4
 Best Local Similarity: 66.67% Mismatches: 0
 Query Match: 76.19% Indels: 0
 DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x AAK15273 (1-148)

QY 1 HisAenHISAenHISAenHISAenHISAen 12
 Db 11 CACAACCAACCAACCAACCAACCAACCAACCAAC 46

RESULT 81

ABS40587
 ID ABS40587 standard; DNA; 148 BP.

XX	AB540587;
AC	
XX	
DT	25-FEB-2003 (first entry)
XX	
DE	Human liver single exon probe, SEQ ID No 15577.
XX	
KW	Human; single exon nucleic acid probe; liver; cirrhosis;
KM	hyperlipoproteinemia; hyperlipidaemia; hypercholesterolaemia;
OS	coronary heart disease; ss.
XX	
PN	Homo sapiens.
XX	
PD	WO200157273-A2.
XX	
PF	09-AUG-2001.
XX	
PR	30-JAN-2001; 2001WO-US000664.
XX	
PR	04-FEB-2000; 2000US-0180312P.
XX	26-MAY-2000; 2000US-0207456P.
PR	30-JUN-2000; 2000US-00608408.
XX	
PR	03-AUG-2000; 2000US-00632366.
XX	21-SEP-2000; 2000US-0234687P.
PR	27-SEP-2000; 2000US-0236359P.
XX	04-OCT-2000; 2000GB-00024263.
PA	(MOLE-) MOLECULAR DYNAMICS INC.
PI	
PT	Penn SG, Hanzel DK, Chen W, Rank DR,
DR	WP1; 2001-488698/53.
XX	
PT	Human genome-derived single exon nucleic acid probes useful for analyzing
PT	gene expression in human adult liver.
XX	
PS	Claim 4; SEQ ID NO 15577; 658pp; English.
XX	
CC	The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC	measuring human gene expression in a sample derived from human adult
CC	liver, comprising one of 13109 defined nucleotide sequences given in the
CC	specification (or complements/ fragments). The probe hybridises at high
CC	stringency to a nucleic acid molecule expressed in the human adult liver.
CC	(I) may be used for predicting, measuring and displaying gene expression
CC	in samples derived from human adult liver. The genes identified may be
CC	involved in genetic liver diseases such as cirrhosis,
CC	hyperlipoproteinemia, hyperlipidaemia and hypercholesterolaemia which is
CC	associated with coronary heart disease. ABS25011-ABS51005 represent human
CC	liver single exon nucleic acid probes of the invention. Note: The
CC	sequence information for this patent does not appear in the printed
CC	specification but was obtained in electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 148 BP; 60 A; 56 C; 14 G; 18 T; 0 U; 0 Other;
XX	
Alignment Scores:	
Pred. No.:	1.19 Length: 148
Score:	64.00 Matches: 8
Percent Similarity:	100.00% Conservative: 4
Best Local Similarity:	66.67% Mismatches: 0
Query Match:	76.19% Indels: 0
DB:	4 Gaps: 0
US-09-858-332G-15 (1-12) x AB540587 (1-148)	
OY	1 HisAenhi sAsnhi sAsnhi sAsnhi sAsnhi sAsn 12
Db	: :: :: :: ::
	11 CACAACCACCAACCAACCAACCAACCAACCAACCAAC 46
AA107454	
AA107454 standard; DNA; 148 BP.	
XX	

AC AA107454;
XX
DT 09-OCT-2001 (first entry)
XX
DE Probe #7445 used to measure gene expression in human breast sample.
XX
KW Probe; human; breast disease; breast cancer; development disorder; ss;
inflammatory disease; proliferative breast disease; non-carcinoma tumour
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US000661.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-0060840P.
PR 03-AUG-2000; 2000US-0063236P.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX MPI; 2001-476286/51.
DR novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
XX
PS Claim 25; SEQ ID NO 7445; 322pp; English.

XX The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and non-
CC carcinoma tumors. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 148 BP; 60 A; 56 C; 14 G; 18 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1	Length: 148
Score: 64.00	Matches: 8
Percent Similarity: 100.00%	Conservative: 4
Best Local Similarity: 66.67%	Mismatches: 0
Query Match: 76.19%	Indels: 0
DB: 5	Gaps: 0

US-09-858-332G-15 (1-12) x AA107454 (1-148)

CY 1 HHSAGNHLSAENHLSAENHLSAENHLSAENHLSAEN 12
||| | ||| | ||| : ||| : ||| : ||| : ||| :
Db 11 CACAACCACCAACCAACCAACCAACCAACCAACCAAC 46

RESULT 83

ID ABSI14969 standard; DNA; 148 BP.

XX ABSI14969;

XX DT 19-AUG-2002 (first entry)

[illegible]

```

XX AC ABA43823;
XX DT 01-FEB-2002 (first entry)
XX DE Human breast cell single exon nucleic acid probe #2518.
XX KW Human; microarray; single exon probe; gene expression; breast; disease;
XX cancer; 88.
XX OS Homo sapiens.
XX MN W0200157271-A2.
XX PD 09-AUG-2001.
XX PF 30-JUN-2001; 2001MO-US000662.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 07-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PS WPI; 2001-496933/54.
XX PT New spatially-addressable set of single exon nucleic acid probes, useful
XX PT for measuring gene expression in sample derived from human breast,
XX PT comprises number of single exon nucleic acid probes.
XX PS Claim 1; SEQ ID NO 2518; 327bp + Sequence Listing; English.
XX CC The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human breast and BT 474 cells. The method involves contacting the
XX CC probes with a collection of detectably labelled nucleic acids derived
XX CC from mRNA of human breast, and then measuring the label bound to each
XX CC probe of the microarray. The probes are useful for verifying the
XX CC expression of regions of genomic DNA predicted to encode proteins. They
XX CC are useful for gene discovery, and for determining predisposition and/or
XX CC prognosing breast disease. Gene expression analysis is useful for
XX CC assessing the toxicity of chemical agents on cells. The microarray of
XX CC this invention presents a far greater diversity of probes for measuring
XX CC gene expression, with far less bias than expressed sequence tag
XX CC microarrays. The method is suitable for rapid production of functional
XX CC information from genomic sequence. The present sequence is a single exon
XX CC nucleic acid probe of the invention. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIP0 at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 381 BP; 131 A; 95 C; 72 G; 83 T; 0 U; 0 Other;

Alignment Scores:
Score: 3.09 Length: 381
Pred. No.: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 4 Gaps: 0

US-09-8558-332G-15 (1-12) x ABA43823 (1-381)

```

RESULT 89
AB24034
ID ABA24034 standard; DNA; 381 BP.
XX
AC ABA24034;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #2500 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-48899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX hearts.
XX
PS Claim 1; SEQ ID NO 2500; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 381 BP; 131 A; 95 C; 72 G; 83 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 3.09 Length: 381
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 4 Gaps: 0
US-09-858-332G-15 (1-12) x ABA24034 (1-381)
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
DB 274 CACACCAACCAACCAACCAACCAACCAACCAACCAAC 309
RESULT 90
AAK27997
ID AAK27997 standard; DNA; 381 BP.
XX

AC AAK27997;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 2554.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 2554; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 381 BP; 131 A; 95 C; 72 G; 83 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 3.09 Length: 381
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 4 Gaps: 0
US-09-858-332G-15 (1-12) x AAK27997 (1-381)
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
DB 274 CACACCAACCAACCAACCAACCAACCAACCAACCAAC 309
RESULT 91
AAK2558
ID AAK02558 standard; DNA; 381 BP.
XX
AC AAK02558;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 2549.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX ss.
XX

```

OS Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX
XX PF 30-JAN-2001; 2001WO-US000667.
XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
XX Example 4; SEQ ID NO 2549; 650bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancer. The present sequence is one of the probes of the
XX invention
XX
XX SO Sequence 381 BP; 131 A; 95 C; 72 G; 83 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.09 Length: 381
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x AAK02558 (1-381)

QY 1 Hisaenhisaenhisaenhisaenhisaenhisa 12
Db 274 CACAACCAACCAACCAACCAACCAACCAACCAAC 309

RESULT 92
ABS27588
ID ABS27588 standard; DNA; 381 BP.
XX
XX AC ABS27588;
XX
XX 25-FEB-2003 (first entry)
XX
XX Human liver single exon probe, SEQ ID NO 2578.
XX
XX DE
XX
XX Human; single exon nucleic acid probe; liver; cirrhosis;
XX hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX coronary heart disease; ss.
XX
XX OS Homo sapiens.
XX
XX WO200157273-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US000664.
XX

```

```

PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-48898/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
XX Claim 1; SEQ ID NO 2578; 658bp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENp) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX CC specification (or complements/ fragments). The probe hybridises at high
XX CC stringency to a nucleic acid molecule expressed in the human adult liver.
XX CC (I) may be used for predicting, measuring and displaying gene expression
XX CC in samples derived from human adult liver. The genes identified may be
XX CC involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX CC associated with coronary heart disease. ABS25011-ABS51005 represent human
XX CC liver single exon nucleic acid probes of the invention. Note: The
XX CC sequence information for this patent does not appear in the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences

SO Sequence 381 BP; 131 A; 95 C; 72 G; 83 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.09 Length: 381
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x ABS27588 (1-381)

QY 1 Hisaenhisaenhisaenhisaenhisaenhisa 12
Db 274 CACAACCAACCAACCAACCAACCAACCAACCAAC 309

RESULT 93
AA102485
ID AA102485 standard; DNA; 381 BP.
XX
XX AC AA102485;
XX
XX 09-OCT-2001 (first entry)
XX
XX Probe #2476 used to measure gene expression in human breast sample.
XX
XX DE
XX
XX Probe; human; breast disease; breast cancer; development disorder; ss;
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
XX OS Homo sapiens.
XX
XX WO200157270-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 29-JAN-2001; 2001WO-US000661.
XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX

```


PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-476286/51.
 XX
 PT Novel single exon nucleic acid probe used to measuring gene expression in
 XX a human breast.
 PS
 XX Claim 25; SEQ ID NO 2476; 322pp; English.
 CC The present invention relates to novel single exon nucleic acid probes.
 CC The present sequence is one such probe. The probes are useful for
 CC measuring human gene expression in a human breast sample, where the probe
 CC hybridizes at high stringency to a nucleic acid expressed in the human
 CC breast. The probes are useful for predicting, diagnosing, grading,
 CC staging, monitoring and prognosing diseases of the human breast,
 CC particularly those diseases with polygenic aetiology. The diseases
 CC include: breast cancer, disorders of development, inflammatory diseases
 CC of the breast, fibrocystic changes, proliferative breast disease and non-
 CC carcinoma tumours. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 381 BP; 131 A; 95 C; 72 G; 83 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3.09 Length: 381
 Score: 64.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 4
 Best Local Similarity: 66.67% Mismatches: 0
 Query Match: 76.19% Indels: 0
 DB: 5 Gaps: 0
 US-09-858-332G-15 (1-12) x AA102485 (1-381)
 QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
 |||||
 ID ABS02467 standard; DNA; 381 BP.
 XX
 AC ABS02467;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human genome-derived single exon probe from lung SEQ ID No 2458.
 XX
 XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KM chronic obstructive pulmonary disease; interstitial lung disease;
 KM familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KM Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KM primary ciliary dyskinesia; pulmonary hypertension;
 KM hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000665.

XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 XX measure gene expression in human lung samples.
 PS
 XX Claim 1; SEQ ID NO 2458; 634pp; English.
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC; the novel set of probes which hybridize at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a single exon probe of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 381 BP; 131 A; 95 C; 72 G; 83 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3.09 Length: 381
 Score: 64.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 4
 Best Local Similarity: 66.67% Mismatches: 0
 Query Match: 76.19% Indels: 0
 DB: 6 Gaps: 0
 US-09-858-332G-15 (1-12) x ABS02467 (1-381)
 QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
 |||||

XX Sequence 479 BP; 76 A; 38 C; 227 G; 138 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 3.89 Length: 479
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x AA110862 (1-479)

QY 1 H1sAsnH1sAsnH1sAsnH1sAsnH1sAsn 12
ID ABA52513 standard; DNA; 479 BP.
Db 457 CACCACCATTAACCAACCAACCAACCAACCAACCACTAAC 422

RESULT 97
ABA52513/C
ID ABA52513 standard; DNA; 479 BP.
XX
AC ABA52513;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #818.
XX
KM Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN W0200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 1; SEQ ID NO 818; 639pp + Sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ffp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 479 BP; 76 A; 38 C; 227 G; 138 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 3.89 Length: 479
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0

DB: 4 Gaps: 0
US-09-858-332G-15 (1-12) x ABA52513 (1-479)

QY 1 H1sAsnH1sAsnH1sAsnH1sAsnH1sAsn 12
ID ABA52513 standard; DNA; 479 BP.
Db 457 CACCACCATTAACCAACCAACCAACCAACCAACCACTAAC 422

RESULT 98
AA132122/C
ID AA132122 standard; DNA; 479 BP.
XX
AC AA132122;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #808 used to measure gene expression in human placenta sample.
XX
KM Probe; microarray; human; placenta; antenatal diagnosis;
KM genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN W0200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488937/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 808; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 479 BP; 76 A; 38 C; 227 G; 138 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 3.89 Length: 479
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x AA132122 (1-479)

QY 1 H1sAsnH1sAsnH1sAsnH1sAsnH1sAsn 12
ID ABA42091/C
Db 457 CACCACCATTAACCAACCAACCAACCAACCAACCACTAAC 422

RESULT 99
ABA42091/C
ID ABA42091 standard; DNA; 479 BP.

```

XX AC ABA42091;
XX DT 01-FEB-2002 (first entry)
XX DE Human breast cell single exon nucleic acid probe #786.
XX KW Human; microarray; single exon probe; gene expression; breast; disease;
XX KM cancer; ss.
XX OS Homo sapiens.
XX PN WO200157271-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000662.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-496933/54.
XX PT New spatially-addressable set of single exon nucleic acid probes, useful
XX PT for measuring gene expression in sample derived from human breast,
XX PT comprises number of single exon nucleic acid probes.
XX PS Claim 1; SEQ ID NO 766; 327pp + Sequence Listing; English.
XX PS
XX CC The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human breast and BT 474 cells. The method involves contacting the
XX CC probes with a collection of detectably labelled nucleic acids derived
XX CC from mRNA of human breast, and then measuring the label bound to each
XX CC probe of the microarray. The probes are useful for verifying the
XX CC expression of regions of genomic DNA predicted to encode proteins. They
XX CC are useful for gene discovery, and for determining predisposition and/or
XX CC prognosing breast disease. Gene expression analysis is useful for
XX CC assessing the toxicity of chemical agents on cells. The microarray of
XX CC this invention presents a far greater diversity of probes for measuring
XX CC gene expression, with far less bias than expressed sequence tag
XX CC microarrays. The method is suitable for rapid production of functional
XX CC information from genomic sequence. The present sequence is a single exon
XX CC nucleic acid probe of the invention. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 479 BP; 76 A; 38 C; 227 G; 138 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 3.89 Length: 479
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 4 Gaps: 0
US-09-858-332G-15 (1-12) x ABA42091 (1-479)
QY 1 HisAenH:AsnH:AsnH:AsnH:AsnH:Asn 12
Db 457 CACCAACATACCAACCAACCAACCAACCAACCAATAC 422

```

```

RESULT 100
ABA22302/c
ID ABA22302 standard; DNA; 479 BP.
XX AC ABA22302;
XX DT 23-JAN-2002 (first entry)
XX DE Probe #768 for gene expression analysis in human heart cell sample.
XX KW Human; gene expression; heart; microarray; vascular system; probe;
XX KM cardiovascular disease; hypertension; cardiac arrhythmia;
XX KM congenital heart disease; ss.
XX OS Homo sapiens.
XX PN WO200157274-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000666.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488899/53.
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT hearts.
XX PS Claim 1; SEQ ID NO 768; 530pp; English.
XX PS
XX CC The present invention relates to single exon nucleic acid probes for
XX CC measuring human gene expression in a sample derived from human heart. The
XX CC present sequence is one such probe. The probes may be used for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from the human heart via microarrays. By measuring gene expression, the
XX CC probes are useful for predicting, diagnosing, grading, staging,
XX CC monitoring and prognosing diseases of the human heart and vascular system
XX CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX CC congenital heart disease. Note: The sequence data for this patent did not
XX CC form part of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 479 BP; 76 A; 38 C; 227 G; 138 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 3.89 Length: 479
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 4 Gaps: 0
US-09-858-332G-15 (1-12) x ABA22302 (1-479)
QY 1 HisAenH:AsnH:AsnH:AsnH:AsnH:Asn 12
Db 457 CACCAACATACCAACCAACCAACCAACCAACCAATAC 422
Search completed: October 4, 2005, 11:58:33
Job time : 498 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: October 4, 2005, 11:47:07 ; Search time 3110 Seconds
(without alignments)
146.872 Million cell updates/sec

Title: US-09-858-332G-15
Perfect score: 84
Sequence: 1 NNNNNNNNNNN 12

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xld
-O=/cgmr2.1/USPTO.spool.p/US0985832/runat_04102005_122257_4305/app_query.fasta_1.199
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=100
-DOCALLGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=100 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0985832 @CGN 1.1 5180 @runat_04102005_122257_4305 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
C 1	84	100.0	125	4	BJ340611 BJ340611
C 2	84	100.0	174	4	BJ429957 BJ429957
C 3	84	100.0	209	1	AU053723 AU053723
C 4	84	100.0	214	5	BQ586973 BQ586973
C 5	84	100.0	264	1	AU053106 AU053106
C 6	84	100.0	298	2	BF547803 BF547803
C 7	84	100.0	317	7	DA7535 DA7535
C 8	84	100.0	318	9	CE715092 CE715092
C 9	84	100.0	319	1	AU061961 AU061961

C 10	84	100.0	413	4	BG669342 BG669342
C 11	84	100.0	417	1	AU034405 AU034405
C 12	84	100.0	426	1	AU008498 AU008498
C 13	84	100.0	427	1	AU033706 AU033706
C 14	84	100.0	427	1	AU033893 AU033893
C 15	84	100.0	446	1	AU043889 AU043889
C 16	84	100.0	453	1	AU040068 AU040068
C 17	84	100.0	459	1	AU036977 AU036977
C 18	84	100.0	466	6	CB714934 CB714934
C 19	84	100.0	467	4	BJ431306 BJ431306
C 20	84	100.0	472	4	BJ429835 BJ429835
C 21	84	100.0	476	4	BJ431249 BJ431249
C 22	84	100.0	480	9	CR294373 CR294373
C 23	84	100.0	488	8	AZ321550 AZ321550
C 24	84	100.0	494	4	BJ388347 BJ388347
C 25	84	100.0	496	6	C90140 C90140
C 26	84	100.0	504	7	CK618125 CK618125
C 27	84	100.0	508	6	CB264620 CB264620
C 28	84	100.0	533	8	AQ491972 AQ491972
C 29	84	100.0	538	6	C22962 C22962
C 30	84	100.0	540	1	AA874794 AA874794
C 31	84	100.0	546	4	BJ429095 BJ429095
C 32	84	100.0	556	4	BJ430064 BJ430064
C 33	84	100.0	559	4	BG375827 BG375827
C 34	84	100.0	560	1	AU265802 AU265802
C 35	84	100.0	560	4	BJ329834 BJ329834
C 36	84	100.0	571	4	BJ430176 BJ430176
C 37	84	100.0	571	4	BJ432285 BJ432285
C 38	84	100.0	577	5	BUR86048 BUR86048
C 39	84	100.0	582	4	BJ369186 BJ369186
C 40	84	100.0	583	6	CB586337 CB586337
C 41	84	100.0	593	4	BJ429471 BJ429471
C 42	84	100.0	604	4	BJ433687 BJ433687
C 43	84	100.0	606	4	BJ429190 BJ429190
C 44	84	100.0	609	2	BF237103 BF237103
C 45	84	100.0	609	4	BJ433986 BJ433986
C 46	84	100.0	609	5	BQ783159 BQ783159
C 47	84	100.0	611	4	BJ432773 BJ432773
C 48	84	100.0	616	4	BJ414802 BJ414802
C 49	84	100.0	616	4	BJ428216 BJ428216
C 50	84	100.0	618	4	BJ433790 BJ433790
C 51	84	100.0	620	4	BJ431653 BJ431653
C 52	84	100.0	622	4	BJ401757 BJ401757
C 53	84	100.0	622	4	BJ433898 BJ433898
C 54	84	100.0	624	4	BJ430817 BJ430817
C 55	84	100.0	625	4	BM778956 BM778956
C 56	84	100.0	626	4	BJ428793 BJ428793
C 57	84	100.0	639	4	BJ432833 BJ432833
C 58	84	100.0	644	9	CR320264 CR320264
C 59	84	100.0	652	4	BI271416 BI271416
C 60	84	100.0	652	4	BJ432931 BJ432931
C 61	84	100.0	657	4	BJ431681 BJ431681
C 62	84	100.0	658	6	CD373395 CD373395
C 63	84	100.0	661	5	BQ781099 BQ781099
C 64	84	100.0	662	4	BI281721 BI281721
C 65	84	100.0	663	1	AV682385 AV682385
C 66	84	100.0	663	4	BJ429328 BJ429328
C 67	84	100.0	667	4	BJ433069 BJ433069
C 68	84	100.0	668	4	BJ430907 BJ430907
C 69	84	100.0	676	4	BJ434653 BJ434653
C 70	84	100.0	678	4	BJ431437 BJ431437
C 71	84	100.0	677	7	CV521375 CV521375
C 72	84	100.0	677	4	BJ429677 BJ429677
C 73	84	100.0	678	4	BG448016 BG448016
C 74	84	100.0	679	4	BJ431232 BJ431232
C 75	84	100.0	690	4	BJ430517 BJ430517
C 76	84	100.0	695	9	CG931836 CG931836
C 77	84	100.0	697	4	BJ434010 BJ434010
C 78	84	100.0	697	7	CO389214 CO389214
C 79	84	100.0	701	4	BJ432382 BJ432382
C 80	84	100.0	714	8	AQ874103 AQ874103
C 81	84	100.0	717	4	BJ430955 BJ430955
C 82	84	100.0	718	4	BJ445670 BJ445670

C 83 84 100.0 723 4 BJA29204 BJA29204
C 84 84 100.0 724 4 BJA34636 BJA34636
C 85 84 100.0 725 4 BJA29573 BJA29573
C 86 84 100.0 725 7 CN834352 AGENCOURT
C 87 84 100.0 731 4 BJA36240 BJA36240
C 88 84 100.0 735 9 CR277961 Forward S
C 89 84 100.0 736 2 BF385371
C 90 84 100.0 736 6 C93043
C 91 84 100.0 738 7 CO387676
C 92 84 100.0 741 4 BJA01045
C 93 84 100.0 749 9 BX211625
C 94 84 100.0 759 9 AG542244 Mus muscu
C 95 84 100.0 760 9 CO427845 UI-M-HW0-
C 96 84 100.0 761 7 BJA30651 BJA30651
C 97 84 100.0 762 4 BJA31956 BJA31956
C 98 84 100.0 768 4 BJA32448 BJA32448
C 99 84 100.0 774 4 BJA31981 BJA31981
C 100 84 100.0 774 4 BJA31981 BJA31981

ALIGNMENTS

RESULT 1
LOCUS BJA340611/c 125 bp mRNA linear EST 07-MAR-2002
DEFINITION BJA340611 Dictyostelium discoideum cDNA library, AF Dictyostelium
ACCESSION BJA340611
VERSION BJA340611.1 GI:19248973
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE 1 (bases 1 to 125)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T.
TITLE Full length cDNA of Dictyostelium discoideum at the aggregation
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadaeu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
source
1..125
Location/Qualifiers
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dda3c02"
/sex="mat A"
/dev_stage="Aggregation stage"
/clone_lib="Dictyostelium discoideum cDNA library, AF"

ORIGIN

Alignment Scores:

Pred. No.: 0.019 Length: 125
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BJA340611 (1-125)

QY 1 Hisaenhisaenhisaenhisaenhisaen 12

DB 121 CACATCACAATCACAATCACAATCACAATCACAAT 86

RESULT 2
BJA29957/c

LOCUS BJA29957 174 bp mRNA linear EST 13-MAR-2002
DEFINITION BJA29957 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone dda5d17 3', mRNA sequence.
ACCESSION BJA29957
VERSION BJA29957.1 GI:19404679
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE 1 (bases 1 to 174)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadaeu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
source
1..174
Location/Qualifiers
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv5d17"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Alignment Scores:
Pred. No.: 0.0268 Length: 174
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BJA29957 (1-174)

QY 1 Hisaenhisaenhisaenhisaenhisaen 12

DB 127 CACATCACAATCACAATCACAATCACAATCACAAT 92

RESULT 3
AUG53723

LOCUS AUG53723 209 bp mRNA linear EST 28-APR-1999
DEFINITION AUG53723 Dictyostelium discoideum SL (H.Urushihara) Dictyostelium
discoideum cDNA clone SLJ555, mRNA sequence.
ACCESSION AUG53723
VERSION AUG53723.1 GI:4702205
KEYWORDS EST.

SOURCE

Dictyostelium discoideum
Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

REFERENCE

1 (bases 1 to 209)
Mori,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mitra,B.N., Pi,M., Sato,S., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
developmental cDNA in Dictyostelium discoideum

JOURNAL

Unpublished (1998)
Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
Tel: 81-298-53-4664
Fax: 81-298-53-6614

PROJECT = Dictyostelium discoideum cDNA project in Japan.
Location/Qualifiers

FEATURES

source
1. .209
/organism="Dictyostellium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SLJ555"
/dev_stage="slug"
/clone_1ib="Dictyostellium discoideum SL (H.Urushihara)"

ORIGIN
Alignment Scores:
Pred. No.: 0.0324 Length: 209
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Beet Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-858-332g-15 (1-12) x AU053723 (1-209)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
|||||
2 CACACCAACCAACCAACCAACCAACCAAC 37

RESULT 4
B0586973 214 bp mRNA linear EST 06-DEC-2002
LOCUS B0586973
DEFINITION B012352-024-011-P08-SP6 MP1Z-ADIS-024-leaf Beta vulgaris cDNA clone
024-011-P08 5-PRIME, mRNA sequence.
B0586973
VERSION B0586973.1 GI:26116555
KEYWORDS EST.
SOURCE Beta vulgaris
ORGANISM Beta vulgaris
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Amaranthaceae; Beta.
1 (bases 1 to 214)
Herwig,R., Schulz,B., Weishaar,B., Hennig,S., Steinfach,M.,
Drungowski,M., Stahl,D., Wruck,W., Menze,A., O'Brien,J., Lehrsch,H.
and Radelof,U.
Construction of a 'unigene' cDNA clone set by oligonucleotide
fingerprinting allows access to 25 000 potential sugar beet genes
Plant J. 32 (5), 845-857 (2002)
22362189
PUBMED 12472698
COMMENT Contact: Weishaar B
ADIS DNA core facility at MP1Z
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weishaar@mp1z-koeln.mpg.de
Insert Length: 214 Std Error: 0.00
Plate: 11 row: P column: 08
Seq primer: SP6; CATACGATTAGTGTCACACTATAG.
Location/Qualifiers
1. .214
/organism="Beta vulgaris"
/mol_type="mRNA"
/cultiivar="KWS2320 (double haploid, monogerm breeding
line)"
/db_xref="GABI:185888"
/db_xref="taxon:161934"
/clone="024-011-P08"
/tissue_type="leaf"
/lab_host="EMDH10B"
/clone_1ib="MP1Z-ADIS-024-leaf"
/note="Vector: pCMVSPORT6; Site 1: SalI; Site 2: NotI;
cDNA library from sugar beet. Library provided by KWS
Kleinanzelebener Saatgut AG Einbeck, Germany, contact:
b.schulze@kws.de; cloning sites SalI-NotI, primer sites and
orientation:
SP6-SalI-CCACGCGTCG-5prime-cDNA-polyA-CC-NotI-TI; Note:

Sequencing granted in the context of the GABI-Beet
project, local PI: Dr. Katharina Schneider, coordinator:
Prof. Christian Jung; Sequence submission managed by
RZPD/GABI-Primary database: <http://gabi.rzpd.de>

ORIGIN
Alignment Scores:
Pred. No.: 0.0332 Length: 214
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Beet Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-858-332g-15 (1-12) x B0586973 (1-214)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
|||||
169 CACACCAACCAACCAACCAACCAACCAAC 204

RESULT 5
AU053106 264 bp mRNA linear EST 28-APR-1999
LOCUS AU053106
DEFINITION AU053106 Dictyostellium discoideum SL (H.Urushihara) Dictyostellium
discoideum cDNA clone SLF703, mRNA sequence.
AU053106
VERSION AU053106.1 GI:4701589
KEYWORDS EST.
SOURCE Dictyostellium discoideum
ORGANISM Dictyostellium discoideum
Eukaryota; Mycetozoa; Dictyostelidia; Dictyostellium.
1 (bases 1 to 264)
REFERENCE Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
Developmental cDNA in Dictyostellium discoideum
Unpublished (1998)
TITLE Contact: Hideko Urushihara
JOURNAL Institute of Biological Sciences
COMMENT University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
Tel: 81-298-53-4664
Fax: 81-298-53-6614
Email: hideko@biol.tsukuba.ac.jp
PROJECT = Dictyostellium discoideum cDNA project in Japan.
Location/Qualifiers
1. .264
/organism="Dictyostellium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SLF703"
/dev_stage="slug"
/clone_1ib="Dictyostellium discoideum SL (H.Urushihara)"

FEATURES
source
1. .264
/organism="Beta vulgaris"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SLF703"
/dev_stage="slug"
/clone_1ib="Dictyostellium discoideum SL (H.Urushihara)"

ORIGIN
Alignment Scores:
Pred. No.: 0.0414 Length: 264
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Beet Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-858-332g-15 (1-12) x AU053106 (1-264)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
|||||
41 CACACCAACCAACCAACCAACCAACCAAC 76

RESULT 6
BF547803 298 bp mRNA linear EST 11-DEC-2000
LOCUS BF547803

DEFINITION UI-R-EO-bw-f-10-0-UI.r1 UI-R-EO Rattus norvegicus cDNA clone
 UI-R-EO-bw-f-10-0-UI 5', mRNA sequence.
 ACCESSION BF547803
 VERSION BF547803.1 GI:11638910
 KEYWORDS EST
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 298)
 AUTHORS Ronaldo M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 CDNA Library Preparation: M.B. Soares Lab Clone distribution:
 Clones will be available through Research Genetics (www.resgen.com)
 This clone is also available through the I.M.A.G.E. Consortium at
 LNL (info@image.lnl.gov). IMAGE ID=1769930 The following
 repetitive elements were found in this cDNA sequence: 227-279,
 >[CAAn]Simple Repeat
 Seq primer: M13 Forward.
 FEATURES
 source Location/Qualifiers
 1..298
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-EO-bw-f-10-0-UI"
 /dev_stage="embryonic"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-EO"
 /note="Vector: p773D-Pac (Pharmacia) with a modified
 polylinker; Site 1: NotI; Site 2: EcoRI; This library
 consists of a mixture of individually tagged normalized
 libraries constructed from 8, 12 and 18-day embryo. The
 tag is a string of 3-5 nucleotides present between the
 Not I site and the oligo-dT track which allows
 identification of the library of origin of a clone within
 the mixture."

Alignment Scores:
 Pred. No.: 0.047 Length: 298
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0
 US-09-858-332g-15 (1-12) x BF547803 (1-298)

QY 1 Hisaenhisaenhisaenhisaenhisaen 12
 Db 239 CACAACCATTAACCAACCAACCAACCAACCAAC 274

RESULT 7 317 bp mRNA linear EST 03-APR-2002
 D47535
 LOCUS R1C513103A Rice green shoot Oryza sativa (japonica cultivar-group)
 DEFINITION CDNA, mRNA sequence.
 ACCESSION D47535
 VERSION D47535.1 GI:701244
 KEYWORDS EST.

SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 317)
 AUTHORS Sasaki, T., Miyao, A. and Yamamoto, K.
 TITLE Rice cDNA from callus 1995
 JOURNAL Unpublished (1995)
 COMMENT Contact: Takuji Sasaki
 National Institute of Agrobiological Resources
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
 305-8602, Japan
 Tel: 81-298-38-7441
 Fax: 81-298-38-7468
 Email: tsasaaki@abdr.affrc.go.jp, URL: http://rsgp.dna.affrc.go.jp/
 PROJECT = "RGP".
 FEATURES
 source Location/Qualifiers
 1..317
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone_lib="Rice green shoot"
 /note="Green shoot (8 days old)"

ORIGIN
 Alignment Scores:
 Pred. No.: 0.0501 Length: 317
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 7 Gaps: 0
 US-09-858-332g-15 (1-12) x D47535 (1-317)

QY 1 Hisaenhisaenhisaenhisaenhisaen 12
 Db 151 CACAACCAACCAACCAACCAACCAACCAACAT 186

RESULT 8 318 bp DNA linear GSS 29-SEP-2003
 LOCUS CE715092/c
 DEFINITION tigr-gss-dog-17000369508467 Dog Library Canis familiaris genomic,
 genomic survey sequence.
 ACCESSION CE715092
 VERSION CE715092.1 GI:37034530
 KEYWORDS GSS.
 SOURCE Canis familiaris (dog)
 ORGANISM Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 REFERENCE 1 (bases 1 to 318)
 AUTHORS Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
 Rusch, D.B., Delcher, A.L., Pop, W., Wang, W., Frazer, C.M. and
 Venter, J.C.
 TITLE The dog genome: survey sequencing and comparative analysis
 JOURNAL Science 301 (5641), 1898-1903 (2003)
 MEDLINE 22875432
 PUBMED 14512627
 COMMENT Contact: Kirkness EF
 The Institute for Genomic Research
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-0200
 Fax: 301-838-0208
 Email: ekirknes@tigr.org
 Class: shotgun.
 FEATURES
 source Location/Qualifiers
 1..318
 /organism="Canis familiaris"
 /mol_type="genomic DNA"

ORIGIN

/strain="Standard Poodie"
/db_xref="taxon:9615"
/clone_1ib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from peripheral blood"

Alignment Scores:

Pred. No.:	0.0503	Length:	318
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-858-332g-15 (1-12) x CE715092 (1-318)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
|||||
254 CATATCATATCATATCATATCATATCATATCATAT 219

RESULT 9

AU061961 319 bp mRNA linear EST 20-MAY-1999
LOCUS AU061961 Dictyostelium discoidium SL (H.Urushihara) Dictyostelium
DEFINITION discoidium cDNA clone SLG781, mRNA sequence.
ACCESSION AU061961
VERSION AU061961.1 GI:4883065
KEYWORDS EST.
SOURCE Dictyostelium discoidium
ORGANISM Dictyostelium discoidium
Bukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.

REFERENCE 1 (bases 1 to 319)
Mori, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M., Yoshino, R., Mitera, B.N., Pi, M., Sato, T., Takemoto, K., Yasukawa, H., Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.,
Developmental cDNA in Dictyostelium discoidium
Unpublished (1998)
CONTACT: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
Tel: 81-298-53-4664
Fax: 81-298-53-6614
Email: hideko@biol.tsukuba.ac.jp

PROJECT = Dictyostelium discoidium cDNA project in Japan.
FEATURES

source

1. .319
Location/Qualifiers
/organism="Dictyostelium discoidium"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SLG781"
/dev_stage="slug"
/clone_1ib="Dictyostelium discoidium SL (H.Urushihara)"

ORIGIN

Alignment Scores:

Pred. No.:	0.0505	Length:	319
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-858-332g-15 (1-12) x AU061961 (1-319)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
|||||
85 CACACCAACCAACCAACCAACCAACCAACCAAC 120

RESULT 10

BG669342

LOCUS BG669342 413 bp mRNA linear EST 30-APR-2001
DEFINITION DRNAED03 Rat DRG library Rattus norvegicus cDNA clone DRNAED03
5', mRNA sequence.
ACCESSION BG669342
VERSION BG669342.1 GI:13891264
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 413)
Xiao, H.S., Huang, Q.H., Zhang, F.X., Bao, L., Lu, Y.J., Guo, C., Yang, L., Huang, W.J., Fu, G., Xu, S.H., Cheng, X.P., Yan, Q., Zhu, Z.D., Zhang, X., Chen, Z., Han, Z.G. and Zhang, X.
Identification of gene expression profile of dorsal root ganglion in the rat peripheral axotomy model of neuropathic pain
Proc. Natl. Acad. Sci. U.S.A. 99 (12), 8360-8366 (2002)

TITLE JOURNAL MEDLINE PUBMED
22056133
12060780
COMMENT Contact: Zhang Xu
Laboratory of Sensory System
Institute of Neuroscience
320 Yue Yang Road, Shanghai 200031, P.R.China
Tel: 86-21-64748700-121
Fax: 86-21-64713446
Email: xu.zhang@ion.ac.cn

This clone is also available at Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong New Area, P.R.China. Please contact with Zhang Xu (xu.zhang@ion.ac.cn) or Han Zeguang (hanzg@chgc.sh.cn)
PCR PRIMERS
FORWARD: T3
BACKWARD: T7
Seq primer: T3
POLYA=No.

FEATURES

source

1. .413
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="DRNAED03"
/sex="male"
/tissue_type="dorsal root ganglion"
/dev_stage="adult"
/clone_1ib="Rat DRG library"

ORIGIN

Alignment Scores:

Pred. No.:	0.0661	Length:	413
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-858-332g-15 (1-12) x BG669342 (1-413)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
|||||
213 CACACCATATACCAACCAACCAACCAACCAAC 248

RESULT 11

AU034405 417 bp mRNA linear EST 28-APR-1999
LOCUS AU034405 Dictyostelium discoidium SL (H.Urushihara) Dictyostelium
DEFINITION discoidium cDNA clone SLG829, mRNA sequence.
ACCESSION AU034405
VERSION AU034405.1 GI:3799829
KEYWORDS EST.
SOURCE Dictyostelium discoidium
ORGANISM Dictyostelium discoidium

REFERENCE 1 (bases 1 to 417)
AUTHORS Mori, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M., Yoshino, R., Mita, B.N., Pl.M., Sato, T., Takemoto, K., Yasukawa, H., Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.
TITLE The Dictyostelium developmental cDNA project: generation and analysis of expressed sequence tags from the first-finger stage of development
JOURNAL DNA Res. 5 (6), 335-340 (1998)
MEDLINE 9916227
PUBMED 10048482
COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
Tel: 81-298-53-4664
Fax: 81-298-53-6614
Email: hideko@biol.tsukuba.ac.jp
PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
POLYA=No.

FEATURES
source Location/Qualifiers
1..417
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SLC829"
/dev_stage="slug"
/clone_lib="Dictyostelium discoideum SL (H.Urushihara)"

ORIGIN
Alignment Scores:
Pred. No.: 0.0668 Length: 417
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-858-332g-15 (1-12) x AU034405 (1-417)

QY 1 Hisaenhisaenhisaenhisaenhisaenhisa 12
Db 129 CACAATCACAATCACAATCACAATCACAATCACAAT 164

RESULT 12
LOCUS AI008498 426 bp mRNA linear EST 25-JAN-1999
DEFINITION EST202949 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone
REMARKS 3' end, mRNA sequence.
ACCESSION AI008498
VERSION AI008498
KEYWORDS EST
SOURCE Rattus sp.
Rattus sp.
Rattus sp.
ORGANISM Rattus sp.
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Rattus.
1 (bases 1 to 426)
Lee, N.H., Gloddek, A., Chandra, I., Mason, T.M., Quackenbush, J.,
Kerlavages, A.R. and Adams, M.D.
Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat
Gene Index
Unpublished (1998)
JOURNAL Contact: Lee, NH
COMMENT The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@igr.org
Seq primer: M3-21.
Location/Qualifiers
1..426
source

/organism="Rattus sp."
/mol_type="mRNA"
/db_xref="ATCC (inhost):2016683"
/db_xref="taxon:10118"
/clone="RMA109"
/dev_stage="embryo 8, 12, 18 dpc"
/clone_lib="Normalized rat embryo, Bento Soares"
/note="Vector: pT7T3Pac; Site_1: EcoRI; Site_2: NotI"

ORIGIN
Alignment Scores:
Pred. No.: 0.0683 Length: 426
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-858-332g-15 (1-12) x AI008498 (1-426)

QY 1 Hisaenhisaenhisaenhisaenhisaenhisa 12
Db 322 CACAACATTAACCAACCAACCAACCAACCAAC 287

RESULT 13
LOCUS AU033706 427 bp mRNA linear EST 28-APR-1999
DEFINITION AU033706 Dictyostelium discoideum SL (H.Urushihara) Dictyostelium
discoideum cDNA clone SLB340, mRNA sequence.
ACCESSION AU033706
VERSION AU033706
KEYWORDS EST
SOURCE AU033706.1 GI:3799130
ORGANISM Dictyostelium discoideum
Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
REFERENCE 1 (bases 1 to 427)
Mori, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M.,
Yoshino, R., Mita, B.N., Pl.M., Sato, T., Takemoto, K., Yasukawa, H.,
Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.
The Dictyostelium developmental cDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development
JOURNAL DNA Res. 5 (6), 335-340 (1998)
MEDLINE 9916227
PUBMED 10048482
COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
Tel: 81-298-53-4664
Fax: 81-298-53-6614
Email: hideko@biol.tsukuba.ac.jp
PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
POLYA=No.

FEATURES
source Location/Qualifiers
1..427
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SLB340"
/dev_stage="slug"
/clone_lib="Dictyostelium discoideum SL (H.Urushihara)"

ORIGIN
Alignment Scores:
Pred. No.: 0.0684 Length: 427
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-858-332g-15 (1-12) x AU033706 (1-427)

QY 1 HisAnHISAnHISAnHISAnHISAnHISAn 12
|||||

Db 207 CACAATCACAAATCACAAATCACAAATCACAAAT 242

RESULT 14
LOCUS AU033893 427 bp mRNA linear EST 28-APR-1999
DEFINITION AU033893 Dictyostelium discoideum SL (H.Urushihara) Dictyostelium
discoideum cDNA clone SLB578, mRNA sequence.

ACCESSION AU033893
VERSION AU033893.1 GI:3799317
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

REFERENCE 1 (bases 1 to 427)
Mori, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M.,
Yoshino, R., Mitra, B.N., Pi, M., Sato, T., Takemoto, K., Yasukawa, H.,
Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.
The Dictyostelium developmental cDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development

TITLE DNA Res. 5 (6), 335-340 (1998)

JOURNAL MEDLINE 99156227
PUBMED 10048482

COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
Tel: 81-298-53-4664
Fax: 81-298-53-6614
Email: hideko@iol.tsukuba.ac.jp
PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
POLYA=No.

FEATURES
source 1. 427
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SLB578"
/dev_stage="slug"
/clone_1lb="Dictyostelium discoideum SL (H.Urushihara)"

ORIGIN

Alignment Scores:
Pred. No.: 0.0684 Length: 427
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-858-332g-15 (1-12) x AU033893 (1-427)

QY 1 HisAnHISAnHISAnHISAnHISAnHISAn 12
|||||

Db 207 CACAATCACAAATCACAAATCACAAATCACAAAT 242

RESULT 15
LOCUS BU434889 446 bp mRNA linear EST 13-MAR-2002
DEFINITION BU434889 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone dv25f07 3', mRNA sequence.

ACCESSION BU434889
VERSION BU434889.1 GI:19409611
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

REFERENCE 1 (bases 1 to 446)

AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadenu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@gene.nig.ac.jp.

FEATURES
source 1. 446
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dv25f07"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1lb="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Alignment Scores:
Pred. No.: 0.0716 Length: 446
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BU434889 (1-446)

QY 1 HisAnHISAnHISAnHISAnHISAnHISAn 12
|||||

Db 149 CACAATCACAAATCACAAATCACAAATCACAAAT 114

RESULT 16
LOCUS AU040068 453 bp mRNA linear EST 29-MAR-1999
DEFINITION AU040068 Dictyostelium discoideum SL (H.Urushihara) Dictyostelium
discoideum cDNA clone SLA355, mRNA sequence.

ACCESSION AU040068
VERSION AU040068.1 GI:4009308
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

REFERENCE 1 (bases 1 to 453)
Mori, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M.,
Yoshino, R., Mitra, B.N., Pi, M., Sato, T., Takemoto, K., Yasukawa, H.,
Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.
The Dictyostelium developmental cDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development

TITLE DNA Res. 5 (6), 335-340 (1998)

JOURNAL MEDLINE 99156227
PUBMED 10048482

COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
Tel: 81-298-53-4664
Fax: 81-298-53-6614
Email: hideko@iol.tsukuba.ac.jp
PROJECT = 'Dictyostelium discoideum cDNA project in Japan'.
Location/Qualifiers

FEATURES
source 1. 453
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SLA355"

ORIGIN /dev_stage="slug"
/clone_lib="Dictyostelium discoideum SL (H.Urushihara)"

Alignment Scores:

Pred. No.:	0.0728	Length:	453
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-858-332G-15 (1-12) x AU040068 (1-453)

QY 1 Hisaenhi:saenhi:saenhi:saenhi:saenhi:saen 12
DB 233 CACATCACATCACATCACATCACATCACATCACAT 268

RESULT 17

AU036977 459 bp mRNA linear EST 29-MAR-1999
LOCUS AU036977 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
DEFINITION AU036977 dictyostelium cDNA clone SSB896, mRNA sequence.
ACCESSION AU036977
VERSION AU036977.1 GI:3983730
KEYWORDS EST.

SOURCE

Dictyostelium discoideum
Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
1 (bases 1 to 459)

REFERENCE

Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochida,H. and Tanaka,Y.
The Dictyostelium developmental cDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development

TITLE

DNA Res. 5 (6), 335-340 (1998)

JOURNAL

MEDLINE

PUBMED

COMMENT

Institute of Biological Sciences
University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
Tel: 81-298-53-4664
Fax: 81-298-53-6614
Email: hideko@biol.tsukuba.ac.jp

PROJECT = 'Dictyostelium discoideum cDNA project in Japan'.
Location/Qualifiers

FEATURES

source

1..459
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SSB896"
/dev_stage="slug"
/clone_lib="Dictyostelium discoideum SS (H.Urushihara)"

ORIGIN

Alignment Scores:

Pred. No.:	0.0738	Length:	459
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-858-332G-15 (1-12) x AU036977 (1-459)

QY 1 Hisaenhi:saenhi:saenhi:saenhi:saenhi:saen 12

DB 237 CACATCACATCACATCACATCACATCACATCACAT 272

RESULT 18

CB714934

LOCUS CB714934 466 bp mRNA linear EST 10-APR-2003
DEFINITION AMGNNUC:MRPE3-00062-C4-A placenta embryo D17 (10379) Rattus
norvegicus cDNA clone mrpe3-00062-c4 5', mRNA sequence.

ACCESSION CB714934
VERSION CB714934.1 GI:29772082
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Rodentia; Scuriognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 466)

REFERENCE

Amgen EST Program.
Amgen Rat EST Program
Unpublished (2003)
CONTACT: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00062 row: C column: 4.

COMMENT

FEATURES

source

1..466
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="mrpe3-00062-c4"
/tissue_type="placenta embryo"
/clone_lib="placenta embryo D17 (10379)"
/note="Vector: pSPOR1; placenta embryo D17"

ORIGIN

Alignment Scores:

Pred. No.:	0.075	Length:	466
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-858-332G-15 (1-12) x CB714934 (1-466)

QY 1 Hisaenhi:saenhi:saenhi:saenhi:saenhi:saen 12

DB 405 CACAAACCAACCAACCAACCAACCAACCAACCAAC 440

RESULT 19

LOCUS

BJ431306 467 bp mRNA linear EST 13-MAR-2002
DEFINITION BJ431306 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone dv13013 3', mRNA sequence.

ACCESSION BJ431306
VERSION BJ431306.1 GI:19406028
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

REFERENCE

1 (bases 1 to 467)
Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
Full length cDNA of Dictyostelium discoideum at the vegetative
stage
Unpublished (2002)

JOURNAL

Contact: Tadaa Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tehni@genes.nig.ac.jp.

FEATURES

source

1..467
/organism="Dictyostelium discoideum"
/mol_type="mRNA"

ORIGIN
Alignment Scores:
Pred. No.: 0.0752 Length: 467
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv13013"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1ib="Dictyostelium discoideum cDNA library, VF"

US-09-858-332g-15 (1-12) x BJ431306 (1-467)

QY 1 HisA^{snH}isA^{snH}isA^{snH}isA^{snH}isA^{snH}isA^{snH} 12
DB 211 CACATCACAATCACAATCACAATCACAATCACAAT 176

RESULT 20 BJ429835 472 bp mRNA linear EST 13-MAR-2002
LOCUS BJ429835/c Dictyostelium discoideum cDNA library, VF Dictyostelium
DEFINITION dictyostelium cDNA clone ddv5d06 3', mRNA sequence.
ACCESSION BJ429835
VERSION BJ429835.1 GI:19404557
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.

REFERENCE 1 (bases 1 to 472)
AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage

JOURNAL Unpublished (2002)
COMMENT Contact: Tadau Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..472
Location/Qualifiers

/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv5d06"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1ib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Alignment Scores:
Pred. No.: 0.076 Length: 472
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BJ429835 (1-472)

QY 1 HisA^{snH}isA^{snH}isA^{snH}isA^{snH}isA^{snH}isA^{snH} 12
DB 161 CACATCACAATCACAATCACAATCACAATCACAAT 126

RESULT 21
BJ431249/c

LOCUS BJ431249 476 bp mRNA linear EST 13-MAR-2002
DEFINITION BJ431249 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone ddv13007 3', mRNA sequence.
ACCESSION BJ431249
VERSION BJ431249.1 GI:19405971
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.

REFERENCE 1 (bases 1 to 476)
AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage

JOURNAL Unpublished (2002)
COMMENT Contact: Tadau Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..476
Location/Qualifiers

/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv13007"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1ib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Alignment Scores:
Pred. No.: 0.0767 Length: 476
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BJ431249 (1-476)

QY 1 HisA^{snH}isA^{snH}isA^{snH}isA^{snH}isA^{snH}isA^{snH} 12
DB 210 CACATCACAATCACAATCACAATCACAATCACAAT 175

RESULT 22 CR294373 480 bp DNA linear GSS 28-FEB-2004
LOCUS CR294373/c Medicago truncatula BAC ends cultivar Jemalong A17 of Medicago
DEFINITION truncatula, genomic survey sequence.
ACCESSION CR294373
VERSION CR294373.1 GI:44684797
KEYWORDS GSS.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.

REFERENCE 1 (bases 1 to 480)
AUTHORS Direct Submision
TITLE Submitted (25-FEB-2004) Genoscope - Centre National de Sequencage ;
BP 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
JOURNAL Location/Qualifiers

FEATURES
source
1..480
Location/Qualifiers
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/cultivar="Jemalong A17"
/db_xref="taxon:3880"

RESULT 25
C90140 496 bp mRNA linear EST 20-APR-1998
LOCUS C90140 Dictyostelium discoideum SS (H Urushihara) Dictyostelium
DEFINITION dictyostelium cDNA clone SS1136, mRNA sequence.
ACCESSION C90140
VERSION C90140.1 GI:3059760
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Maeda,B.N., Pi,M., Saito,T., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
DEVELOPMENTAL CDNA in Dictyostelium discoideum
JOURNAL Unpublished (1998)
COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
Tel: 81-298-53-4664
Fax: 81-298-53-6614
Email: hideko@biol.tsukuba.ac.jp.
Location/Qualifiers
FEATURES
source
1. .496
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strat="AX4"
/db_xref="taxon:44689"
/clone="SS1136"
/dev_stage="slug"
/clone_1lb="Dictyostelium discoideum SS (H. Urushihara)"
ORIGIN
Alignment Scores:
Pred. No.: 0.08 Length: 496
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-858-332g-15 (1-12) x C90140 (1-496)
QY 1 H1aSnH1aSnH1aSnH1aSnH1aSnH1aSn 12
DB 201 CACAAATCAATCAATCAATCAATCAATCAAT 236
RESULT 26
CK618125 504 bp mRNA linear EST 26-JAN-2004
LOCUS CK618125 m105h06 5', mRNA sequence.
DEFINITION Dictyostelium mus musculus cDNA clone
CK618125
VERSION CK618125.1 GI:41339011
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Ida,H., Boylan,S., Weigel,A., Smit-McBride,Z., Chao,A., Gao,J.,
Buchhoff,P., Wietow,G. and Hjelmeland,L.
Expressed sequence tag analysis of mouse retina
JOURNAL Unpublished (2004)
COMMENT Contact: Wietow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov

Plate: 05 row: h column: 06
Seq primer: M13Rpi reverse primer (AB1).
FEATURES
source
1. .504
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strat="C57B16J"
/db_xref="taxon:10090"
/clone="m105h06"
/sex="male"
/tissue_type="retina"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_1lb="mouse retina, unamplified; mk/ml"
/note="Organ: Eye; Vector: pSPORT1; 270ug total RNA was
extracted from 200 adult male mouse retinas. A
directionally cloned cDNA library in the pSPORT1
vector (Life Technologies) was constructed at Bioserve
Biotechnology (Laurel MD) essentially following the
protocol of the Superscript Plasmid System full details
of which are contained in the manufacturer's instruction
manual (http://www.lifetech.com/). First strand synthesis
was carried out using a Not I primer-adaptor
[5'-GCACTAGTCTAGATCGGAGCGGCCGCTT)15-3']. cDNA was
cloned in Not I/Sal I sites. EST analysis was performed on
the unamplified library at the NIH Intramural Sequencing
Center (NISC)."
ORIGIN
Alignment Scores:
Pred. No.: 0.0814 Length: 504
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-858-332g-15 (1-12) x CK618125 (1-504)
QY 1 H1aSnH1aSnH1aSnH1aSnH1aSnH1aSn 12
DB 109 CACAAACCAACCAACCAACCAACCAACCAAC 144
RESULT 27
CB264620 528 bp mRNA linear EST 06-NOV-2003
LOCUS CB264620 50-B014659-035-002-C14-T7R MP1Z-ADIS-035 Arabidopsis thaliana cDNA
DEFINITION clone MP1Zp2000C142Q 5-PRIME, mRNA sequence.
ACCESSION CB264620
VERSION CB264620.1 GI:32889393
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS Schmidt,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altman,T.,
Mitchell-Olds,T. and Weisshaar,B.
Large-scale identification and analysis of genome-wide
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
JOURNAL Genome Res. 13 (6), 1250-1257 (2003)
MEDLINE 22683290
PUBMED 12799357
COMMENT Contact: Weisshaar B
ADIS DNA core facility at MP1Z
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weisshaar@ipz-koeln.mpg.de
Insert Length: 528 Std Error: 0.00
Plate: 2 row: C column: 14
Seq primer: T7R; CTAATACGACTCACTATAGGA.

FEATURES
Source

Location/Qualifiers
1..528
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/culivar="Achkarren-2 (Ak-2)"
/db_xref="GABI:594788"
/db_xref="taxon:3702"
/clone="MP12P200C142Q"
/issue_type="inflorescence"
/lab_host="E. coli TOP10"
/clone_lib="MP12-ADIS-035"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; cDNA library from Arabidopsis thaliana, accession Achkarren-2; inflorescences from flower buds to young siliques; library was made at the Max-Planck-Institute for Plant Breeding Research, Cologne, Germany; cloning sites SalI-NotI, primer sites and orientation:
T7-SalI-CCACGCTCCG-5prime-cDNA-polyA-CC-NotI-SP6; Note: Sequencing granted in the context of the GABI Arabidopsis Bundel: Genetic Diversity, 'Establishment of high-efficiency SNP-based mapping tools and development of methods for genome-wide mutation detection' PI: Bernd Weisshaar; Sequence submission managed by RZPD/GABI-Primary database; <http://gabi.rzpd.de> This clone is available from RZPD; contact RZPD (clone@rzpd.de) for further information."

ORIGIN

Alignment Scores:

Pred. No.: 0.0854 Length: 528
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-858-332g-15 (1-12) x CB264620 (1-528)

QY 1 HisAenHisAenHisAenHisAenHisAenHisAen 12

Db 98 CACAATCACAATCACAATCACAATCACAATCACAAT 63

RESULT 28

AQ491972

LOCUS AQ491972 533 bp DNA linear GSS 24-APR-1999

DEFINITION V100810 mtn-3xHA/lacZ insertion library Saccharomyces cerevisiae

ACCESSION AQ491972

VERSION AQ491972.1 GI:4677846

KEYWORDS GSS.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anjay.kumar@yale.edu
Sequence corresponds to the extreme 5' end of mtn-3xHA/lacZ (ending in TCAGACCCC) followed by native Saccharomyces cerevisiae genomic sequence immediately upstream, generated by single read using sequencing primer indicated above; enables determination of site of

mtn-3xHA/lacZ insertion.
Seq primer: GGCTCTTCTTCTTGAGAGTAC
Class: transposon-tagged.
Location/Qualifiers
1..533

FEATURES

Source

/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"
/lab_host="E. coli"
/clone_lib="mtn-3xHA/lacZ insertion library"
/note="Vector: pHS56-SalI; A yeast genomic DNA library (lacking mitochondrial DNA) was prepared in pHS56-SalI; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mtn-3xHA/lacZ multitransposon containing lacZ, URA3, and tet resistance."

ORIGIN

Alignment Scores:

Pred. No.: 0.0863 Length: 533
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-09-858-332g-15 (1-12) x AQ491972 (1-533)

QY 1 HisAenHisAenHisAenHisAenHisAenHisAen 12

Db 109 CATATCATATCATATCATATCATATCATATCATAT 144

RESULT 29

C22962

LOCUS C22962 538 bp mRNA linear EST 02-APR-1997

DEFINITION C22962 Dictyostelium discoideum FC (H.Urushihara) Dictyostelium

ACCESSION C22962

VERSION C22962.1 GI:1922111

KEYWORDS EST.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

Source

Location/Qualifiers
1..538
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="KAX3"
/db_xref="taxon:44689"
/clone="FC-AL02"
/sex="mat A"
/issue_type="gamete"
/dev stage="sexual"
/clone_lib="Dictyostelium discoideum FC (H. Urushihara)"

ORIGIN

Alignment Scores:

Pred. No.: 0.0871 Length: 538
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-858-332g-15 (1-12) x C22962 (1-538)

Qy 1 H1sA8nH1sA8nH1sA8nH1sA8nH1sA8n 12
 |||||
 Db 460 CACACCAACCAACCAACCAACCAACCAAC 495

RESULT 30
 AA874794/c 540 bp mRNA linear EST 04-JUL-1999
 LOCUS AA874794
 DEFINITION UI-R-E0-bw-f-10-0-UI-s1 UI-R-E0 Rattus norvegicus cDNA clone
 UI-R-E0-bw-f-10-0-UI 3' similar to gi1189378|gb|M38188|HUMOGC Human
 unknown protein from clone pHR74 mRNA, complete cde, mRNA
 sequence.
 AA874794
 ACCESSION AA874794.1 GI:4230677
 VERSION AA874794.1
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 540)
 REFERENCE Bernaldo,M.F., Lennon,G. and Soares,M.B.
 AUTHORS Normalization and subtraction: two approaches to facilitate gene
 TITLE discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT On Mar 20, 1998 this sequence version replaced gi:2979742.
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 The sequence tag present in the cDNA between the NotI site and the
 oligo-dT track served to identify it as a clone from the normalized
 adult 8-Day-Embryo library. cDNA Library Preparation: M. Fatima
 Bernaldo, Ph.D. Clone distribution: clones will be available through
 Research Genetics This clone is also available through the
 I.M.A.G.E. Consortium at LBNL (info@image.llnl.gov). IMAGE
 ID=169930
 Seq primer: M13 Forward
 POLYA=No.

FEATURES
 source
 Location/Qualifiers
 1..540
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-E0-bw-f-10-0-UI"
 /dev_stage="embryonic"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-E0"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: NotI; Site 2: EcoRI; This library
 consists of a mixture of individually tagged normalized
 libraries constructed from 8, 12 and 18-day embryo. The
 tag is a string of 3-5 nucleotides present between the
 Not I site and the oligo-dT track which allows
 identification of the library of origin of a clone within
 the mixture."

ORIGIN
 Alignment Scores:
 Pred. No.: 0.0875 length: 540
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0

DB: 1 Gaps: 0

US-09-858-332g-15 (1-12) x AA874794 (1-540)

Qy 1 H1sA8nH1sA8nH1sA8nH1sA8nH1sA8n 12
 |||||
 Db 537 CATAACCAACCAACCAACCAACCAACCAACAT 502

RESULT 31
 BU429095/c 546 bp mRNA linear EST 13-MAR-2002
 LOCUS BU429095
 DEFINITION BU429095 Dictyostelium discoideum cDNA library, VF Dictyostelium
 discoideum cDNA clone ddtv2107 3', mRNA sequence.
 BU429095
 ACCESSION BU429095.1 GI:19404786
 VERSION BU429095.1
 KEYWORDS Dictyostelium discoideum
 SOURCE Dictyostelium discoideum
 ORGANISM Dictyostelium discoideum
 Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 1 (bases 1 to 546)
 REFERENCE Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
 AUTHORS Full length cDNA of Dictyostelium discoideum at the vegetative
 TITLE stage
 JOURNAL unpublished (2002)
 COMMENT Contact: Tadaeu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.

FEATURES
 source
 Location/Qualifiers
 1..546
 /organism="Dictyostelium discoideum"
 /mol_type="mRNA"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="ddv2107"
 /sex="mat A"
 /dev_stage="Growth phase"
 /clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN
 Alignment Scores:
 Pred. No.: 0.0885 length: 546
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BU429095 (1-546)

Qy 1 H1sA8nH1sA8nH1sA8nH1sA8nH1sA8n 12
 |||||
 Db 166 CACAATCAATCAATCAATCAATCAATCAATCAAT 131

RESULT 32
 BU430064/c 556 bp mRNA linear EST 13-MAR-2002
 LOCUS BU430064
 DEFINITION BU430064 Dictyostelium discoideum cDNA library, VF Dictyostelium
 discoideum cDNA clone ddtv5023 3', mRNA sequence.
 BU430064
 ACCESSION BU430064.1 GI:19404786
 VERSION BU430064.1
 KEYWORDS Dictyostelium discoideum
 SOURCE Dictyostelium discoideum
 ORGANISM Dictyostelium discoideum
 Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 1 (bases 1 to 556)
 REFERENCE Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
 AUTHORS Full length cDNA of Dictyostelium discoideum at the vegetative
 TITLE stage
 JOURNAL unpublished (2002)

COMMENT

Contact: Tadao Shin-ji
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@gene.nig.ac.jp.

FEATURES

SOURCE

Location/Qualifiers
1. .556
/organism="Dictyostellium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv5023"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostellium discoideum cDNA library, VF"

ORIGIN

Alignment Scores:

Pred. No.: 0.0902 Length: 556
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BG375827 (1-556)

QY 1 HisaAnH1sAAnH1sAAnH1sAAnH1sAAnH1sAAn 12

DB 180 CACAAATCAATCAATCAATCAATCAATCAATCAAT 145

RESULT 33

BG375827/c

LOCUS BG375827 559 bp mRNA linear EST 12-MAR-2001

DEFINITION UI-R-CS0-bcf-h-04-0-UI.s1 UI-R-CS0 Rattus norvegicus cDNA clone

ACCESSION

BG375827

VERSION

BG375827.1

KEYWORDS

EST

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 559)

REFERENCE

Bonaldo, M.F., Lennon, G. and Soares, M.B.

AUTHORS

TITLE

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newson Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: benito-soares@iowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
normalized rat heart pool library cDNA library Preparation: M.B.
Research Genetics (www.resgen.com) The following repetitive
>(CAA)n#5Simple repeat
Seq primer: M13 Forward
POLYA=yes.

Location/Qualifiers

SOURCE

1. .559
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CS0-bcf-h-04-0-UI"
/dev_stage="ADULT"
/lab_host="PH108 (Life Technologies)"
/clone_lib="UI-R-CS0"
/note="Vector: pYT3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CS0
library is a normalized library constructed from the
following rat heart tissues: embryonic day 17, embryonic
day 19, embryonic day 21, adult day 1, adult day 12, adult
day 75, adult day 200. For a detailed description of the
library from which this clone was derived, please visit
our web site at ratat.eng.uiowa.edu. The subtraction has
been previously described in (Bonaldo, Lennon and Soares,
Genome Research 6:791-806, 1996)
TAG TISSUE=rat heart pool
TAG LIB=UI-R-CS0
TAG_SEQ=ATGAATGATAC"

ORIGIN

Alignment Scores:
Pred. No.: 0.0907 Length: 559
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BG375827 (1-559)

QY 1 HisaAnH1sAAnH1sAAnH1sAAnH1sAAnH1sAAn 12

DB 544 CACAACCATTAACCAACCAACCAACCAACCAAC 509

RESULT 34

A2265802

LOCUS

A2265802

DEFINITION

A2265802 VS Dictyostellium discoideum cDNA clone VSR793 5', mRNA

ACCESSION

A2265802

VERSION

A2265802.1

KEYWORDS

EST

SOURCE

ORGANISM

Dictyostellium discoideum

Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.

1 (bases 1 to 560)

REFERENCE

Urushihara, H., Morio, T., Saito, T., Kohara, Y., Koriiki, E., Ochiai, H.,

AUTHORS

Maeda, M., Williams, J.G., Takeuchi, I. and Tanaka, Y.,

TITLE

Analyses of cDNAs from growth and slug stages of Dictyostellium

discoideum

Nucleic Acids Res. 32 (5), 1647-1653 (2004)

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
Tel: 81-298-53-4664
Fax: 81-298-53-6614
Email: hideko@biol.tsukuba.ac.jp.
Location/Qualifiers
1. .560
/organism="Dictyostellium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="VSP793"
/sex="mat A"
/dev_stage="vegetative"
/clone_lib="VS"

ORIGIN

FEATURES

Alignment Scores:

Pred. No.: 0.0909 Length: 560
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-858-332g-15 (1-12) x AU265802 (1-560)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 413 CACATCACATCACATCACATCACATCACATCACAT 448

RESULT 35

BU329834 560 bp mRNA linear EST 05-MAR-2002

LOCUS BU329834 Dictyostelium discoidium cDNA library, AF Dictyostelium

DEFINITION BU329834 Dictyostelium discoidium cDNA clone dda26007 5', mRNA sequence.

ACCESSION BU329834 GI:19159964

KEYWORDS EST. Dictyostelium discoidium

SOURCE Dictyostelium discoidium

ORGANISM Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.

REFERENCE 1 (bases 1 to 560)

AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.

TITLE Full length cDNA of Dictyostelium discoidium at the aggregation

JOURNAL stage

COMMENT Contact: Tadaeu Shin-i

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1. 560

/organism="Dictyostelium discoidium"

/mol_type="mRNA"

/strain="AX4"

/db_xref="taxon:44689"

/clone="dda26007"

/sex="mat A"

/dev_stage="Aggregation stage"

/clone_lib="Dictyostelium discoidium cDNA library, AF"

ORIGIN

Alignment Scores:

Pred. No.: 0.0909 Length: 560
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BU329834 (1-560)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 202 CACATCACATCACATCACATCACATCACATCACAT 237

RESULT 36

BU430176 571 bp mRNA linear EST 13-MAR-2002

LOCUS BU430176 Dictyostelium discoidium cDNA library, VF Dictyostelium

DEFINITION BU430176 Dictyostelium discoidium cDNA clone ddv6h09 3', mRNA sequence.

ACCESSION BU430176 GI:19404898

KEYWORDS EST. Dictyostelium discoidium

SOURCE Dictyostelium discoidium

ORGANISM Dictyostelium discoidium

Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.

REFERENCE 1 (bases 1 to 571)

AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.

TITLE Full length cDNA of Dictyostelium discoidium at the vegetative

JOURNAL stage

COMMENT Contact: Tadaeu Shin-i

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1. 571

/organism="Dictyostelium discoidium"

/mol_type="mRNA"

/strain="AX4"

/db_xref="taxon:44689"

/clone="ddv6h09"

/sex="mat A"

/dev_stage="Growth phase"

/clone_lib="Dictyostelium discoidium cDNA library, VF"

ORIGIN

Alignment Scores:

Pred. No.: 0.0927 Length: 571
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BU430176 (1-571)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 171 CACATCACATCACATCACATCACATCACATCACAT 136

RESULT 37

BU432285 571 bp mRNA linear EST 13-MAR-2002

LOCUS BU432285 Dictyostelium discoidium cDNA library, VF Dictyostelium

DEFINITION BU432285 Dictyostelium discoidium cDNA clone ddv18b10 3', mRNA sequence.

ACCESSION BU432285 GI:19407007

KEYWORDS EST. Dictyostelium discoidium

SOURCE Dictyostelium discoidium

ORGANISM Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.

REFERENCE 1 (bases 1 to 571)

AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.

TITLE Full length cDNA of Dictyostelium discoidium at the vegetative

JOURNAL stage

COMMENT Contact: Tadaeu Shin-i

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1. 571

/organism="Dictyostelium discoidium"

/mol_type="mRNA"

/strain="AX4"

/db_xref="taxon:44689"

/clone="ddv18b10"

/sex="mat A"

/dev_stage="Growth phase"

/clone_lib="Dictyostelium discoidium cDNA library, VF"

ORIGIN

Alignment Scores:

Pred. No.:	0.0927	Length:	571
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-858-332G-15 (1-12) x BU432285 (1-571)

QY 1 HisAenHisAenHisAenHisAenHisAen 12

Db 194 CACAATCACAATCACAATCACAATCACAATCACAAT 159

RESULT 38

BU886048 577 bp mRNA linear EST 17-OCT-2002

LOCUS BU886048/c tremuloides cDNA 5 prime, mRNA sequence.

ACCESSION BU886048.1 GI:24077565

VERSION BU886048.1

KEYWORDS EST.

ORGANISM Populus tremula x Populus tremuloides

Populus tremula x Populus tremuloides

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Malpighiales; Salicaceae; Populus.

1 (bases 1 to 577)

Uneberg, P., Bhalerao, R.R., Jansson, S. and Sterky, F.

The poplar tree transcriptome: Analysis of expressed sequence tags

from multiple libraries

Unpublished (2002)

Contact: BHALERAO RUPALI R.

Umea Plant Science Center

Department of Plant Physiology

University of Umea, 901 87 Umea, Sweden

Tel: +46 90 786 5279

Fax: +46 90 786 6676

Email: rupali.bhalerao@plantphys.umu.se.

Location/Qualifiers

1..577

/organism="Populus tremula x Populus tremuloides"

/mol_type="mRNA"

/db_xref="taxon:47664"

/tissue_type="root"

/clone_lib="Populus root cDNA library"

ORIGIN

Alignment Scores:

Pred. No.:	0.0938	Length:	577
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	5	Gaps:	0

US-09-858-332G-15 (1-12) x BU886048 (1-577)

QY 1 HisAenHisAenHisAenHisAenHisAen 12

Db 272 CATATCATATCATATCATATCATATCATATCATAT 237

RESULT 39

BU369186 582 bp mRNA linear EST 08-MAR-2002

LOCUS BU369186 Dicotyledonum discoidium cDNA library, CF Dicotyledonum

ACCESSION BU369186

VERSION BU369186.1 GI:19278569

KEYWORDS EST.

SOURCE Dicotyledonum discoidium

ORGANISM

Dicotyledonum discoidium

Eukaryota; Mycetozoa; Dicotyledonida; Dicotyledonum.

1 (bases 1 to 582)

Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.

Full length cDNA of Dicotyledonum discoidium at the culmination

stage

Unpublished (2002)

Contact: Tadao Shin-i

Center For Genetic Resource Information

National Institute of Genetics

111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tsunig@genes.nig.ac.jp.

Location/Qualifiers

1..582

/organism="Dicotyledonum discoidium"

/mol_type="mRNA"

/strain="AX4"

/db_xref="taxon:44689"

/clone="ddc49k17"

/sex="mat A"

/dev_stage="Culmination stage"

/clone_lib="Dicotyledonum discoidium cDNA library, CF"

ORIGIN

Alignment Scores:

Pred. No.:	0.0946	Length:	582
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-858-332G-15 (1-12) x BU369186 (1-582)

QY 1 HisAenHisAenHisAenHisAenHisAen 12

Db 284 CACAACCAACCAACCAACCAACCAACCAACCAAC 319

RESULT 40

LOCUS CB586337 583 bp mRNA linear EST 03-APR-2003

DEFINITION AMGNMNC:NRD1-00072-h8-A nrd1 (10855) Rattus norvegicus cDNA clone

ACCESSION CB586337

VERSION CB586337.1 GI:29530378

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 583)

Amgen EST Program

Amgen Rat EST Program

Unpublished (2003)

Contact: Dan Fitzpatrick

Amgen, Inc

One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00072 row, h column: 8.

Location/Qualifiers

1..583

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="nrd1-00072-h8"

/tissue_type="Dorsal Root Ganglia"

/clone_lib="nrd1 (10855)"

/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; rat

ORIGIN

Alignment Scores:
Pred. No.: 0.0948 Length: 583
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-858-332G-15 (1-12) x CBS6337 (1-583)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
|||||
Db 296 CACAACCTAACCACCAACCAACCAACCAAC 331

RESULT 41
LOCUS BJ429471/c 599 bp mRNA linear EST 13-MAR-2002
DEFINITION BJ429471 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone ddv3114 3', mRNA sequence.
ACCESSION BJ429471
VERSION BJ429471.1 GI:19404193
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE 1 (bases 1 to 599)
AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..599
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv3114"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN
Alignment Scores:
Pred. No.: 0.0975 Length: 599
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BJ429471 (1-599)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
|||||
Db 198 CACAATCACATCACATCACATCACATCACAT 163

RESULT 42
LOCUS BJ433687/c 604 bp mRNA linear EST 13-MAR-2002
DEFINITION BJ433687 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone ddv22113 3', mRNA sequence.
ACCESSION BJ433687
VERSION BJ433687.1 GI:19408409
KEYWORDS EST.
SOURCE Dictyostelium discoideum

ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE 1 (bases 1 to 604)
AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..604
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv22113"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN
Alignment Scores:
Pred. No.: 0.0983 Length: 604
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BJ433687 (1-604)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
|||||
Db 158 CACAATCACATCACATCACATCACATCACAT 123

RESULT 43
LOCUS BJ429190/c 606 bp mRNA linear EST 13-MAR-2002
DEFINITION BJ429190 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone ddv2116 3', mRNA sequence.
ACCESSION BJ429190
VERSION BJ429190.1 GI:19403912
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE 1 (bases 1 to 606)
AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..606
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv2116"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Alignment Scores:

Pred. No.: 0.0987 Length: 606
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BU429190 (1-606)

Qy 1 Hisaenhi:saenhi:saenhi:saenhi:saenhi:saen 12

Db 199 CACATCACAATCACAATCACAATCACAATCACAAT 164

RESULT 44

LOCUS

BP237103 609 bp mRNA linear EST 14-NOV-2000
 DEFINITION 602027447F1 NCI CGAP_i19 Mus musculus cDNA clone IMAGE:4162763 5',
 mRNA sequence.

ACCESSION BP237103 GI:11151020

VERSION BP237103.1 GI:11151020

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 609)

AUTHORS NIH-MGC http://mgs.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapds-rc@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINT at:
 http://image.lnl.gov

plate: L1AM9445 row: n column: 12

High quality sequence stop: 609.

Location/Qualifiers

FEATURES

source

1..609

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4162763"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NCI CGAP_i19"

/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.9 kb. Constructed by Life
 Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Alignment Scores:

Pred. No.: 0.0992 Length: 609
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-858-332G-15 (1-12) x BP237103 (1-609)

Qy 1 Hisaenhi:saenhi:saenhi:saenhi:saenhi:saen 12

Db 83 CACACCAACCAACCAACCAACCAACCAACCAAC 118

RESULT 45

BU433986/c

LOCUS

BU433986 609 bp mRNA linear EST 13-MAR-2002
 DEFINITION BU433986 Dictyostelium discoideum cDNA library, VF Dictyostelium
 discoideum cDNA clone d4v23e18 3', mRNA sequence.

ACCESSION

BU433986 GI:19408708

KEYWORDS

EST.

SOURCE Dictyostelium discoideum

ORGANISM

REFERENCE Dictyostelium discoideum
 Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 1 (bases 1 to 609)

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2002)

Contact: Tadao Shin-i

Center For Genetic Resource Information

National Institute of Genetics

111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1..609

/organism="Dictyostelium discoideum"

/mol_type="mRNA"

/strain="AX4"

/db_xref="taxon:44689"

/clone="d4v23e18"

/sex="mat A"

/dev stage="Growth phase"

/clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Alignment Scores:

Pred. No.: 0.0992 Length: 609
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BU433986 (1-609)

Qy 1 Hisaenhi:saenhi:saenhi:saenhi:saenhi:saen 12

Db 152 CACATCACAATCACAATCACAATCACAATCACAAT 117

RESULT 46

LOCUS

BU783159 609 bp mRNA linear EST 26-JUL-2002
 DEFINITION UI-R-PF0-cpl-d-10-0-UI.61 NCI CGAP_PFO Rattus norvegicus cDNA clone
 UI-R-PF0-cpl-d-10-0-UI 3', mRNA sequence.

ACCESSION

BU783159 GI:21991631

KEYWORDS

EST.

SOURCE

ORGANISM

REFERENCE Rattus norvegicus (Norway rat)

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Jeff Stevens
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: DISTRIBUTION: Researchers may obtain clones
from Research Genetics (www.reagen.com).
The following repetitive elements were found in this CDNA
sequence: 504-556 >(CAA)n\$simple_repeat (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source

Location/Qualifiers
1..609
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="UI-R-F0-cpi-d-10-0-UI"
/tissue_type="Mixed tissues"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="NCI CGAP F0"
/note="Vector: pT73-Pac (Pharmacia) with a modified
polylinker; Site 1: Ecor I; Site 2: Not I; UI-R-F0 is a
subtracted CDNA library containing the following
tissue(s): Normal cartilage and SR-JMS Tumor Line. The
subtraction was made according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for these libraries are: CTATGTGACG,
CATCTTGTA.
TAG_TISSUE=rat SRC-JMS tumor line
TAG_LIB=UI-R-F0
TAG_SEQ=CATTCTGTGA"

ORIGIN

Alignment Scores:

Pred. No.: 0.0992 Length: 609
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-858-332g-15 (1-12) x BQ783159 (1-609)

QY 1 H18AsnH18AsnH18AsnH18AsnH18Asn 12

DB 544 CACAACTTAACCAACCAACCAACCAACCAAC 509

RESULT 47
Bj432773/c

LOCUS Bj432773 611 bp mRNA linear EST 13-MAR-2002
DEFINITION dictyostelium discoidium cDNA library, VF Dictyostelium
discoidium cDNA clone ddt19c20 3', mRNA sequence.

ACCESSION Bj432773

VERSION Bj432773.1

KEYWORDS GI:19407495

SOURCE EST.

ORGANISM Dictyostelium discoidium

REFERENCE Dictyostelium discoidium

AUTHORS Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

TITLE 1 (bases 1 to 611)

JOURNAL Unruhara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.

COMMENT Full length cDNA of Dictyostelium discoidium at the vegetative
stage
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856

Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..611
/organism="Dictyostelium discoidium"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv19c20"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoidium cDNA library, VF"

FEATURES

source

Location/Qualifiers
1..611
/organism="Dictyostelium discoidium"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv19c20"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoidium cDNA library, VF"

ORIGIN

Alignment Scores:

Pred. No.: 0.0995 Length: 611
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x Bj432773 (1-611)

QY 1 H18AsnH18AsnH18AsnH18AsnH18Asn 12

DB 163 CACAACTTAACCAACCAACCAACCAACCAAT 128

RESULT 48
Bj414802

LOCUS Bj414802 616 bp mRNA linear EST 10-MAR-2002
DEFINITION dictyostelium discoidium cDNA library, VF Dictyostelium
discoidium cDNA clone ddt20p07 5', mRNA sequence.

ACCESSION Bj414802

VERSION Bj414802.1

KEYWORDS GI:19327313

SOURCE EST.

ORGANISM Dictyostelium discoidium

REFERENCE Dictyostelium discoidium

AUTHORS Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

TITLE 1 (bases 1 to 616)

JOURNAL Unruhara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.

COMMENT Full length cDNA of Dictyostelium discoidium at the vegetative
stage
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES

source

Location/Qualifiers
1..616
/organism="Dictyostelium discoidium"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv20p07"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoidium cDNA library, VF"

ORIGIN

Alignment Scores:

Pred. No.: 0.1 Length: 616
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x Bj414802 (1-616)

QY 1 HisaenHisAenHisAenHisAenHisAen 12
|||||
Db 515 CACAACCAACCAACCAACCAACCAACCAAC 550

RESULT 49
LOCUS BJ428216/c 616 bp mRNA linear EST 11-MAR-2002
DEFINITION BJ428216 Dictyostelium discoidium cDNA library, VF Dictyostelium
discoidium cDNA clone ddv1101 3', mRNA sequence.
ACCESSION BJ428216
VERSION BJ428216.1 GI:19345573
KEYWORDS EST.
SOURCE Dictyostelium discoidium
ORGANISM Dictyostelium discoidium
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE 1 (bases 1 to 616)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T.
TITLE Full length cDNA of Dictyostelium discoidium at the vegetative
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6855
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..616
/organism="Dictyostelium discoidium"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv1101"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1ib="Dictyostelium discoidium cDNA library, VF"

ORIGIN
Alignment Scores:
Pred. No.: 0.1 Length: 616
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BJ428216 (1-616)

QY 1 HisaenHisAenHisAenHisAenHisAen 12
|||||
Db 207 CACAATCACAATCACAATCACAATCACAATCACAAT 172

RESULT 50
LOCUS BJ433790/c 618 bp mRNA linear EST 13-MAR-2002
DEFINITION BJ433790 Dictyostelium discoidium cDNA library, VF Dictyostelium
discoidium cDNA clone ddv22n24 3', mRNA sequence.
ACCESSION BJ433790
VERSION BJ433790.1 GI:19408512
KEYWORDS EST.
SOURCE Dictyostelium discoidium
ORGANISM Dictyostelium discoidium
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE 1 (bases 1 to 618)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T.
TITLE Full length cDNA of Dictyostelium discoidium at the vegetative
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..618
/organism="Dictyostelium discoidium"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv22n24"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1ib="Dictyostelium discoidium cDNA library, VF"

ORIGIN
Alignment Scores:
Pred. No.: 0.101 Length: 618
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BJ433790 (1-618)

QY 1 HisaenHisAenHisAenHisAenHisAen 12
|||||
Db 158 CACAATCACAATCACAATCACAATCACAATCACAAT 123

RESULT 51
LOCUS BJ431653/c 620 bp mRNA linear EST 13-MAR-2002
DEFINITION BJ431653 Dictyostelium discoidium cDNA library, VF Dictyostelium
discoidium cDNA clone ddv15b01 3', mRNA sequence.
ACCESSION BJ431653
VERSION BJ431653.1 GI:19406375
KEYWORDS EST.
SOURCE Dictyostelium discoidium
ORGANISM Dictyostelium discoidium
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE 1 (bases 1 to 620)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T.
TITLE Full length cDNA of Dictyostelium discoidium at the vegetative
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..620
/organism="Dictyostelium discoidium"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv15b01"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1ib="Dictyostelium discoidium cDNA library, VF"

ORIGIN
Alignment Scores:
Pred. No.: 0.101 Length: 620
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BJ431653 (1-620)

Qy 1 H1eAsnH1eAsnH1eAsnH1eAsnH1eAsn 12
|||||
Db 208 CACAATCACAATCACAATCACAATCACAAT 173

RESULT 52
Bj401757 622 bp mRNA linear EST 10-MAR-2002
LOCUS Bj401757 dictyostelium discoidum cDNA library, SF Dictyostelium
DEFINITION discoidum cDNA clone ddel19e12 3', mRNA sequence.
ACCESSION Bj401757
VERSION Bj401757.1 GI:19314674
KEYWORDS EST.
SOURCE Dictyostelium discoidum
ORGANISM Dictyostelium discoidum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE 1 (bases 1 to 622)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE Full length cDNA of Dictyostelium discoidum at the slug stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadau Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES
source
1..622
/organism="Dictyostelium discoidum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddel19e12"
/sex="mat A"
/dev_stage="Slug stage"
/clone_lib="Dictyostelium discoidum cDNA library, SF"

ORIGIN
Alignment Scores:
Pred. No.: 0.101 Length: 622
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x Bj401757 (1-622)

Qy 1 H1eAsnH1eAsnH1eAsnH1eAsnH1eAsn 12
|||||
Db 158 CACAATCACAATCACAATCACAATCACAAT 123

RESULT 53
Bj433898 622 bp mRNA linear EST 13-MAR-2002
LOCUS Bj433898 dictyostelium discoidum cDNA library, VF Dictyostelium
DEFINITION discoidum cDNA clone ddtv23d08 3', mRNA sequence.
ACCESSION Bj433898
VERSION Bj433898.1 GI:19408620
KEYWORDS EST.
SOURCE Dictyostelium discoidum
ORGANISM Dictyostelium discoidum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE 1 (bases 1 to 622)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE Full length cDNA of Dictyostelium discoidum at the vegetative stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadau Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES
source
1..622
/organism="Dictyostelium discoidum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv23d08"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoidum cDNA library, VF"

ORIGIN
Alignment Scores:
Pred. No.: 0.101 Length: 622
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x Bj433898 (1-622)

Qy 1 H1eAsnH1eAsnH1eAsnH1eAsnH1eAsn 12
|||||
Db 163 CACAATCACAATCACAATCACAATCACAAT 128

RESULT 54
Bj340817 624 bp mRNA linear EST 07-MAR-2002
LOCUS Bj340817 dictyostelium discoidum cDNA library, AF Dictyostelium
DEFINITION discoidum cDNA clone dda4c02 3', mRNA sequence.
ACCESSION Bj340817
VERSION Bj340817.1 GI:19249179
KEYWORDS EST.
SOURCE Dictyostelium discoidum
ORGANISM Dictyostelium discoidum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE 1 (bases 1 to 624)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE Full length cDNA of Dictyostelium discoidum at the aggregation stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadau Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES
source
1..624
/organism="Dictyostelium discoidum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dda4c02"
/sex="mat A"
/dev_stage="Aggregation stage"
/clone_lib="Dictyostelium discoidum cDNA library, AF"

ORIGIN
Alignment Scores:
Pred. No.: 0.102 Length: 624
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x Bj340817 (1-624)

Qy 1 Hisaenhisaenhisaenhisaenhisaen 12
 |||||
 Db 167 CACATCAATCAATCAATCAATCAATCAAT 132

RESULT 55
 LOCUS BM778956 625 bp mRNA linear EST 04-MAR-2002
 DEFINITION EST589531 KV2 Medicago truncatula cDNA clone PKV2-20C24, mRNA
 SEQUENCE.
 VERSION BM778956
 ACCESSION BM778956.1 GI:19108624
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
 Medicago.
 1 (bases 1 to 625)
 Vandenbosch, K., Endre, G., Hur, J., Beremand, P., Town, C.D., Van
 Aken, S., Utecherback, T., Cheung, F., Tsai, J. and Fraser, C.M.
 ESTs from roots of Medicago truncatula 48 hr after inoculation with
 Sinorhizobium meliloti
 Unpublished (2002)
 Other ESTs: EST589530
 Contact: Vandenbosch K
 Department of Plant Biology
 University of Minnesota
 220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
 Tel: 612 624 2755
 Fax: 612 625 1738
 Email: kvandenbosch.umn.edu
 TIGR sequence name: MTABH24TKB More information is available at:
 www.medicago.org
 Seq primer: Skmod (CTA GAA CTA gtc gat CC).
 Location/Qualifiers
 1..625
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="genotype A17"
 /db_xref="taxon:3880"
 /clone="PKV2-20C24"
 /issue_type="Seedling roots"
 /dev_stage="2 days post-inoculation with Sinorhizobium
 meliloti"
 /lab_host="E. coli strain SOLR"
 /lab_rib="KV2"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
 XhoI; cDNA was prepared from polyA+ enriched RNA. The
 cDNA was directionally ligated into the UniTap XR vector
 from Stratagene and packaged using Gigapack III Gold
 packaging extracts. Plasmids containing cDNA inserts
 were excised from the recombinant lambda-Zap phage using
 Ex-assist helper phage and propagated in SOLR cells."

ORIGIN

Alignment Scores:
 Pred. No.: 0.102 Length: 625
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BM778956 (1-625)

Qy 1 Hisaenhisaenhisaenhisaenhisaen 12
 |||||
 Db 236 CACATCAATCAATCAATCAATCAATCAAT 271

RESULT 56
 BJA28793/c

LOCUS BJA28793 626 bp mRNA linear EST 13-MAR-2002
 DEFINITION BJA28793 Dictyostelium discoideum cDNA library, VF Dictyostelium
 discoideum cDNA clone ddv1c09 3', mRNA sequence.
 VERSION BJA28793
 ACCESSION BJA28793.1 GI:19403515
 KEYWORDS EST
 SOURCE Dictyostelium discoideum
 ORGANISM Dictyostelium discoideum
 Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
 1 (bases 1 to 626)
 Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
 Full length cDNA of Dictyostelium discoideum at the vegetative
 stage
 Unpublished (2002)
 JOURNAL Contact: Tadao Shin-i
 COMMENT Center For Genetic Resource Information
 National Institute of Genetics
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.
 Location/Qualifiers
 1..626
 /organism="Dictyostelium discoideum"
 /mol_type="mRNA"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="ddv1c09"
 /sex="mat A"
 /dev_stage="Growth phase"
 /clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Alignment Scores:
 Pred. No.: 0.102 Length: 626
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BJA28793 (1-626)

Qy 1 Hisaenhisaenhisaenhisaenhisaen 12
 |||||
 Db 170 CACATCAATCAATCAATCAATCAATCAAT 135

RESULT 57
 BJA32833/c

LOCUS BJA32833 639 bp mRNA linear EST 13-MAR-2002
 DEFINITION BJA32833 Dictyostelium discoideum cDNA library, VF Dictyostelium
 discoideum cDNA clone ddv20b02 3', mRNA sequence.
 VERSION BJA32833
 ACCESSION BJA32833.1 GI:19407555
 KEYWORDS EST.
 SOURCE Dictyostelium discoideum
 ORGANISM Dictyostelium discoideum
 Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
 1 (bases 1 to 639)
 Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
 Full length cDNA of Dictyostelium discoideum at the vegetative
 stage
 Unpublished (2002)
 JOURNAL Contact: Tadao Shin-i
 COMMENT Center For Genetic Resource Information
 National Institute of Genetics
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.
 Location/Qualifiers
 1..639
 /organism="Dictyostelium discoideum"

```

/mo1_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv20b02"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1ib="Dictyostelium discoideum cDNA library, VF"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 0.104 Length: 639
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

```

US-09-858-332G-15 (1-12) x BJ432833 (1-639)

```

Qy 1 HisAenHisAenHisAenHisAenHisAenHisAen 12
    |||||
Db 157 CACAATCATATCATATCATATCATATCATATCATAT 122

```

```

RESULT 58 CR320264 644 bp DNA linear GSS 01-MAR-2004
LOCUS CR320264 Medicago truncatula BAC ends cultivar Jemalong A17 of Medicago
DEFINITION truncatula, genomic survey sequence.
ACCESSION CR320264
VERSION CR320264.1 GI:44866408
KEYWORDS GSS.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 644)
Genoscope.
Direct Submission
Submitted (25-FEB-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr]
Web : www.genoscope.cns.fr
FEATURES
    source
        1..644
            Location/Qualifiers
                /organism="Medicago truncatula"
                /mol_type="genomic DNA"
                /cultivar="Jemalong A17"
                /db_xref="taxon:3880"
                /clone_1ib="MTB1"
                /note="Vector: pIndigoBAC ; Site 1: EcoRI ; Site 2: EcoRI
                ; Debelle F. and Chalmou B.-Genoscope sequence ID :
                mel-46E22RM1"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 0.105 Length: 644
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

```

US-09-858-332G-15 (1-12) x CR320264 (1-644)

```

Qy 1 HisAenHisAenHisAenHisAenHisAenHisAen 12
    |||||
Db 180 CACAATCATATCATATCATATCATATCATATCATAT 215

```

RESULT 59 BI271416 652 bp mRNA linear EST 18-JUL-2001

LOCUS BI271416

DEFINITION NF051D09FLF1078 Developing flower Medicago truncatula cDNA clone

```

ACCESSION NF051D09FL 5', mRNA sequence.
VERSION BI271416
KEYWORDS BI271416.1 GI:14879906
SOURCE EST.
ORGANISM Medicago truncatula (barrel medic)

```

```

REFERENCE 1 (bases 1 to 652)
AUTHORS Torres-Velez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
JOURNAL Medicago truncatula flower library
COMMENT Unpublished (2001)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6692
Email: gdmay@noble.org
Insert Length: 652 Std Error: 0.00
Plate: 051 row: D column: 09
Seq primer: TCACACAGGAACAGCTATGAC.
Location/Qualifiers
1..652
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF051D09FL"
/tissue_type="Developing flowers"
/dev_stage="Developmentally pooled. Contains a mixture of
very young, developing, fully-opened flowers and flowers
in early transition into pods."
/clone_1ib="Developing flower"
/note="Vector: Lambda Zap; cDNA was prepared from polyA+
enriched, pooled samples of equivalent amounts of total
RNA from very young, developing, fully-opened flowers and
flowers transitioning into pods. The cDNA was
directionally ligated into the Uni-Zap XR vector
(Stratagene) and packaged using the Gigapack III Gold
packaging extracts. Phagemids containing cDNA inserts were
in vivo excised from the recombinant Uni-Zap XR vector
using ExAest helper phage and the E. coli strain
Xin-Blue MRF" (Stratagene). Excised plasmids were plated
using SOLR cells."

```

FEATURES

```

source
    1..652
        Location/Qualifiers
            /organism="Medicago truncatula"
            /mol_type="mRNA"
            /db_xref="taxon:3880"
            /clone="NF051D09FL"
            /tissue_type="Developing flowers"
            /dev_stage="Developmentally pooled. Contains a mixture of
            very young, developing, fully-opened flowers and flowers
            in early transition into pods."
            /clone_1ib="Developing flower"
            /note="Vector: Lambda Zap; cDNA was prepared from polyA+
            enriched, pooled samples of equivalent amounts of total
            RNA from very young, developing, fully-opened flowers and
            flowers transitioning into pods. The cDNA was
            directionally ligated into the Uni-Zap XR vector
            (Stratagene) and packaged using the Gigapack III Gold
            packaging extracts. Phagemids containing cDNA inserts were
            in vivo excised from the recombinant Uni-Zap XR vector
            using ExAest helper phage and the E. coli strain
            Xin-Blue MRF" (Stratagene). Excised plasmids were plated
            using SOLR cells."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 0.107 Length: 652
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

```

US-09-858-332G-15 (1-12) x BI271416 (1-652)

```

Qy 1 HisAenHisAenHisAenHisAenHisAenHisAen 12
    |||||
Db 426 CACAATCATATCATATCATATCATATCATATCATAT 461

```

RESULT 60 BJ432931 652 bp mRNA linear EST 13-MAR-2002

LOCUS BJ432931

DEFINITION BJ432931 Dictyostelium discoideum cDNA library, VF Dictyostelium

ACCESSION BJ432931

VERSION BJ432931.1 GI:19407653

KEYWORDS EST.

SOURCE Dictyostelium discoideum

ORIGIN

created for the University of Iowa Program for Rat Gene
Discovery and Mapping (Val Sheffield, Bento Soares and Tom
Casavant)."

Alignment Scores:

Pred. No.:	0.108	Length:	658
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-858-332G-15 (1-12) x CD373395 (1-658)

QY 1 HisAenHisAenHisAenHisAenHisAenHisAen 12
|||||
Db 99 CACACCATRACCAACCAACCAACCAACCAAC 134

RESULT 63

BO781099/c

LOCUS

UI-R-PFO-cpd-j-13-0-UI.s1 NCI_CGAP_PFO Rattus norvegicus CDNA clone

DEFINITION

UI-R-PFO-cpd-j-13-0-UI 3', mRNA sequence.

ACCESSION

BO781099

VERSION

BO781099.1

KEYWORDS

EST.

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 661)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

8889548

CONTACT: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@iowa.edu

Tissue Procurement: Jeff Stevens

CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: DISTRIBUTION: Researchers may obtain clones

from Research Genetics (www.regen.com).

The following repetitive elements were found in this CDNA

sequence: 504-556, >(CAA)n\$Simple_repeat (matched complement)

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES

source

1.661
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="UI-R-PFO-cpd-j-13-0-UI"
/tissue_type="Mixed tissues"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="NCI_CGAP_PFO"
/note="Vector: pT73-Pac (Pharmacia) with a modified
polylinker; Site 1: EcoR I; Site 2: Not I; UI-R-PFO is a
subtracted CDNA library containing the following
tissue(s): Normal cartilage and SR-JWS Tumor line. The
subtraction was made according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. The
oligonucleotide used to prime the synthesis of

ORIGIN

first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for these libraries are: CTATGGACG,
CATCTCTGTA,
TAG_TISSUE=cartilage
TAG_LIB=UI-R-PFO
TAG_SEQ=CTATGGACG"

Alignment Scores:

Pred. No.:	0.108	Length:	661
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	5	Gaps:	0

US-09-858-332G-15 (1-12) x BO781099 (1-661)

QY 1 HisAenHisAenHisAenHisAenHisAenHisAen 12
|||||
Db 544 CACACCATRACCAACCAACCAACCAACCAAC 509

RESULT 64

BI281721/c

LOCUS

UI-R-CTOs-cav-f-10-0-UI.s1 UI-R-CTOs Rattus norvegicus CDNA clone

DEFINITION

UI-R-CTOs-cav-f-10-0-UI 3', mRNA sequence.

ACCESSION

BI281721

VERSION

BI281721.1

KEYWORDS

EST.

SOURCE

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 662)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

8889548

CONTACT: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@iowa.edu

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the CDNA between the NotI site

and the oligo-dT track served to identify it as a clone from the

non-normalized rat brain pool library CDNA library Preparation:

M.B. Soares Lab Clone distribution: clones will be available

FEATURES

source

1.662
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CTOs-cav-f-10-0-UI"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-CTOs"
/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; site 1: Not I; Site 2: Eco RI; The UR-CT05 library is a non-normalized library constructed from the following rat brain tissues: embryonic day 17, embryonic day 19, embryonic day 21, adult day 1, adult day 12, adult day 75, adult day 200. For a detailed description of the library from which this clone was derived, please visit our web site at ratat.eng.uwa.edu. The subcloning has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996).

TAG TISSUE=rat brain pool
TAG LIB=UR-CT05
TAG_SEQ=ACTTC

ORIGIN

Alignment Scores:
Pred. No.: 0.108 Length: 662
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x B1281721 (1-662)

QY 1 HisAenHISAenHISAenHISAenHISAenHISAen 12
Db 544 CACAACTACATCACATCACATCACATCACATCACAT 509

RESULT 65 663 bp mRNA linear EST 01-APR-2002
LOCUS AV823854/RAFL6 Arabidopsis thaliana cDNA clone RAFL6-09-121 5',
DEFINITION mRNA sequence.
ACCESSION AV823854
VERSION AV823854
KEYWORDS EST
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 663)
Oono, Y., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Arakawa, T., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyada, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msek@rkc.riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pBluescript vector as a Set1/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES

Location/Qualifiers
1..663
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL6-09-121"
/dev_stage="plants at various developmental stages from germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAFL6"
/note="Site 1: Set1; Site 2: XhoI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"

ORIGIN

Alignment Scores:
Pred. No.: 0.108 Length: 663
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-858-332g-15 (1-12) x AV823854 (1-663)

QY 1 HisAenHISAenHISAenHISAenHISAenHISAen 12
Db 144 CACATCACATCACATCACATCACATCACATCACAT 109

RESULT 66 663 bp mRNA linear EST 13-MAR-2002
LOCUS BJ429328/Dictyostelium discoidium cDNA library, VF Dictyostelium
DEFINITION BJ429328 Dictyostelium discoidium cDNA clone dv3104 3', mRNA sequence.
ACCESSION BJ429328
VERSION BJ429328
KEYWORDS EST
SOURCE Dictyostelium discoidium
ORGANISM Dictyostelium discoidium
Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
1 (bases 1 to 663)
Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
Full length cDNA of Dictyostelium discoidium at the vegetative stage
Unpublished (2002)
Contact: Tadasi Shin-i
Center for Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp

Journal
COMMENT Unpublished (2002)
Contact: Tadasi Shin-i
Center for Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp

FEATURES

Location/Qualifiers
1..663
/organism="Dictyostelium discoidium"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dv3104"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoidium cDNA library, VF"

ORIGIN

Alignment Scores:
Pred. No.: 0.108 Length: 663
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BJ429328 (1-663)

QY 1 HisAenHISAenHISAenHISAenHISAenHISAen 12
Db 198 CACATCACATCACATCACATCACATCACATCACAT 163

RESULT 67 667 bp mRNA linear EST 13-MAR-2002
LOCUS BJ433069/Dictyostelium discoidium cDNA library, VF Dictyostelium
DEFINITION BJ433069 Dictyostelium discoidium cDNA clone dv20n16 3', mRNA sequence.
ACCESSION BJ433069
VERSION BJ433069
KEYWORDS EST

SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE 1 (bases 1 to 667)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadau Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES
source
1..667
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv20n16"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1ib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Alignment Scores:
Pred. No.: 0.109 Length: 667
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 4
DB: 4

US-09-858-332g-15 (1-12) x BJ433069 (1-667)

Qy 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
|||||
Db 146 CACAAATCACAATCACAATCACAATCACAAT 111

RESULT 68
BJ430907/c 668 bp mRNA linear EST 13-MAR-2002
LOCUS BJ430907 Dictyostelium discoideum cDNA library, VF Dictyostelium
DEFINITION dictyoidem cDNA clone ddv9a09 3', mRNA sequence.
ACCESSION BJ430907
VERSION BJ430907.1 GI:19405629
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadau Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES
source
1..668
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv9a09"
/sex="mat A"
/dev_stage="Growth phase"

ORIGIN /clone_1ib="Dictyostelium discoideum cDNA library, VF"

Alignment Scores:
Pred. No.: 0.109 Length: 668
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 4
DB: 4

US-09-858-332g-15 (1-12) x BJ434653 (1-668)

Qy 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
|||||
Db 152 CACAATCACAATCACAATCACAATCACAAT 117

RESULT 70
BJ431437/c 676 bp mRNA linear EST 13-MAR-2002
LOCUS BJ431437 Dictyostelium discoideum cDNA library, VF Dictyostelium
DEFINITION dictyoidem cDNA clone ddv14m03 3', mRNA sequence.
ACCESSION BJ431437
VERSION BJ431437.1 GI:19406159
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadau Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES
source
1..668
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv24f16"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1ib="Dictyostelium discoideum cDNA library, VF"

KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
AUTHORS 1 (bases 1 to 676)
TITLE Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
Full length cDNA of Dictyostelium discoideum at the vegetative stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tahini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES
SOURCE 1..676
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv14m03"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1lb="Dictyostelium discoideum cDNA library, VF"

ORIGIN
Alignment Scores:
Pred. No.: 0.111 Length: 676
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BJ431437 (1-676)

OY 1 HisAenHisAenHisAenHisAenHisAen 12
|||||
Db 194 CACAAATCAATCAATCAATCAATCAATCAAT 159

RESULT 71
LOCUS CV521375 676 bp mRNA linear EST 06-OCT-2004
DEFINITION 0089P00582.x0.G04 Mimulus guttatus library 2 Mimulus guttatus cDNA
ACCESSION CV521375
VERSION CV521375.1 GI:53847907
KEYWORDS EST.
SOURCE Mimulus guttatus (spotted monkey flower)
ORGANISM Mimulus guttatus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Lamiales; Lamiales incertae sedis; Mimuleae;
Mimulus.
1 (bases 1 to 676)
Willis, J., Vision, T., Dietrich, F.S. and Allen, A.
Mimulus guttatus cDNA sequence
Unpublished (2004)
Contact: Willis J
Department of Biology
Duke University
072-A Biological Sciences Science Drive, Durham, NC 27708, USA
Tel: 919 660 7340
Fax: 919 660 7293
Email: jwillis@duke.edu
Plate: 0089P0058 row: 04 column: G
Seq primer: T7
High quality sequence start: 16
High quality sequence stop: 704.
location/Qualifiers
1..676

FEATURES
SOURCE 1..676

/organism="Mimulus guttatus"
/mol_type="mRNA"
/db_xref="taxon:4155"
/clone="0089P00582.x0.G04"
/clone_1lb="Mimulus guttatus library 2"
/note="Vector: pGEM-T Easy; a Mimulus guttatus cDNA library"

ORIGIN
Alignment Scores:
Pred. No.: 0.111 Length: 676
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BJ429677 (1-676)

OY 1 HisAenHisAenHisAenHisAenHisAen 12
|||||
Db 432 CATATCATATCATATCATATCATATCATATCAAT 467

RESULT 72
LOCUS BJ429677/c 677 bp mRNA linear EST 13-MAR-2002
DEFINITION BJ429677 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone ddv4c17 3', mRNA sequence.
ACCESSION BJ429677.1 GI:19404399
VERSION BJ429677.1
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
REFERENCE 1 (bases 1 to 677)
AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tahini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES
SOURCE 1..677
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv4c17"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1lb="Dictyostelium discoideum cDNA library, VF"

ORIGIN
Alignment Scores:
Pred. No.: 0.111 Length: 677
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BJ429677 (1-677)

OY 1 HisAenHisAenHisAenHisAenHisAen 12
|||||
Db 200 CACAAATCAATCAATCAATCAATCAATCAAT 165

RESULT 73

BG448016 678 bp mRNA linear EST 16-MAR-2001
LOCUS NF104H02BC1F1026 Elicited cell culture Medicago truncatula cDNA
DEFINITION clone NF104H02BC 5', mRNA sequence.
ACCESSION BG448016 GI:13366797
VERSION BG448016.1 GI:13366797
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 678)
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Imman, J.T., Weller, J.W. and May, G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation -
Center for Medicago Genomics Research
Unpublished (2000)
CONTACT: Dixon RA
JOURNAL Plant Biology Division
COMMENT The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 678 Std Error: 0.00
Plate: 104 row: H column: 02
Seq primer: TCACACGAGAAACGCTATGAC.
Location/Qualifiers
1. 678
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF104H02BC"
/tissue_type="Cell cultures derived from root tissues"
/dev_stage="Cell suspensions were subcultured every 14
days. Cells were induced six days after subculture"
/clone_1ib="Elicited cell culture"
/note="Vector: Lambda Zap. Cells were induced with yeast
cell wall extracts equivalent to 50ug/ml glucose in the
final concentration. Samples were taken at 0.5, 1, 12 and
24 hours after induction. Equal amounts of RNA from each
time point were pooled and used for mRNA isolation."

ORIGIN
Alignment Scores:
Pred. No.: 0.111 Length: 678
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-09-858-332G-15 (1-12) x BG448016 (1-678)
QY 1 H15ASNH15ASNH15ASNH15ASNH15ASNH15ASN 12
|||||
Db 264 CACATCATATCATATCATATCATATCATATCATATCATAT 299

RESULT 74
Bj431232 679 bp mRNA linear EST 13-MAR-2002
LOCUS Bj431232 Dictyostelium discoideum cDNA library, VF Dictyostelium
DEFINITION dictyoidem cDNA clone d4v13h09 3', mRNA sequence.
ACCESSION Bj431232 GI:19405954
VERSION Bj431232.1 GI:19405954
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE 1 (bases 1 to 679)
Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-I, T.

TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. 679
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="d4v13h09"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1ib="Dictyostelium discoideum cDNA library, VF"

ORIGIN
Alignment Scores:
Pred. No.: 0.113 Length: 690
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-09-858-332G-15 (1-12) x Bj431232 (1-679)
QY 1 H15ASNH15ASNH15ASNH15ASNH15ASNH15ASN 12
|||||
Db 203 CACATCATATCATATCATATCATATCATATCATATCATAT 168

RESULT 75
Bj430517 690 bp mRNA linear EST 13-MAR-2002
LOCUS Bj430517 Dictyostelium discoideum cDNA library, VF Dictyostelium
DEFINITION dictyoidem cDNA clone d4v7d17 3', mRNA sequence.
ACCESSION Bj430517
VERSION Bj430517.1 GI:19405239
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE 1 (bases 1 to 690)
Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-I, T.
Full length cDNA of Dictyostelium discoideum at the vegetative
stage
Unpublished (2002)
CONTACT: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. 690
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="d4v7d17"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1ib="Dictyostelium discoideum cDNA library, VF"

ORIGIN
Alignment Scores:
Pred. No.: 0.113 Length: 690

Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BJ430517 (1-690)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
 DB 118 CACAATCAATCAATCAATCAATCAATCAATCAAT 83

RESULT 76
 CG931836

LOCUS CG931836 695 bp DNA linear GSS 12-DEC-2003
 DEFINITION MBEL86TR mth2 Medicago truncatula genomic clone 13p4, genomic
 survey sequence.
 ACCESSION CG931836
 VERSION CG931836.1 GI:39797669
 KEYWORDS GSS
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosoids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
 Medicago.

REFERENCE 1 (bases 1 to 695)

AUTHORS Town, C.D., Shetty, D., Koo, H. and Feldjany, T.F.
 TITLE Sequencing of BAC ends from Medicago truncatula
 JOURNAL Unpublished (2003)
 COMMENT Other GSSs: MBEL86TRFB
 CONTACT: Chris Town

9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org

Seq primer: CACGAACAGCTATGACC
 Class: BAC ends.

FEATURES
 source location/Qualifiers

1..695
 /organism="Medicago truncatula"
 /mol_type="genomic DNA"
 /cultivar="genotype A17"
 /db_xref="taxon:3880"
 /clone="13p4"
 /clone_lib="mth2"
 /note="Vector: pBelobAC11, Site_1: HindIII, Site_2:
 HindIII; Cook, D.R. and Kim, D.J, unpublished"

ORIGIN

Alignment Scores:

Pred. No.: 0.114 Length: 695
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-858-332g-15 (1-12) x CG931836 (1-695)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
 DB 177 CACAATCAATCAATCAATCAATCAATCAATCAAT 212

RESULT 77
 BJ434010/c

LOCUS BJ434010 697 bp mRNA linear EST 13-MAR-2002
 DEFINITION Dictyostelium discoideum cDNA library, VF Dictyostelium
 discoideum cDNA clone dv23j14 3', mRNA sequence.
 ACCESSION BJ434010
 VERSION BJ434010.1 GI:19408732
 KEYWORDS EST.

SOURCE

Dictyostelium discoideum

Dictyostelium discoideum

ORGANISM Dictyostelium discoideum

Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

REFERENCE 1 (bases 1 to 697)

AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
 TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
 stage

JOURNAL Unpublished (2002)

CONTACT: Tadao Shin-i

COMMENT Contact: Tadao Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@gene.nig.ac.jp.

FEATURES
 source location/Qualifiers

1..697
 /organism="Dictyostelium discoideum"
 /mol_type="mRNA"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="dv23j14"
 /sex="mat A"
 /dev_stage="Growth phase"
 /clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Alignment Scores:

Pred. No.: 0.114 Length: 697
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BJ434010 (1-697)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
 DB 201 CACAATCAATCAATCAATCAATCAATCAATCAAT 166

RESULT 78
 CG389214

LOCUS CG389214 697 bp mRNA linear EST 30-JUN-2004
 DEFINITION AGENCOURT 26623004 NIH MGC 253 Rattus norvegicus cDNA clone
 IMAGE:7303816 5', mRNA sequence.
 ACCESSION CG389214
 VERSION CG389214.1 GI:49496805
 KEYWORDS EST.

ORGANISM Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 697)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)

CONTACT: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: CGAPDB-r@mail.nih.gov

Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical
 College of Wisconsin

cDNA Library Preparation: Express Genomics.

DNA sequencing by: The I.M.A.G.E. Consortium (LNL)

Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov

Plate: LLM1539 row: 9 column: 14

High quality sequence stop: 666.

Db 337 CATATCATATCATATCATATCATATCATAT 302

RESULT 81
Bj430955/c

LOCUS
DEFINITION Bj430955 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone ddv9j12 3', mRNA sequence.

ACCESSION
Bj430955

VERSION
Bj430955.1 GI:19405677

KEYWORDS
EST.

SOURCE
Dictyostelium discoideum

ORGANISM
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

REFERENCE
1 (bases 1 to 717)
Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
Full length cDNA of Dictyostelium discoideum at the vegetative
stage

JOURNAL
Unpublished (2002)

COMMENT
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1. 717
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv9j12"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Alignment Scores:
Pred. No.: 0.118 Length: 717
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x Bj430955 (1-717)

Qy 1 HisAenHisAenHisAenHisAenHisAenHisAen 12
|||||

Db 155 CACATCATCATCATCATCATCATCATCATCAT 120

RESULT 82
Bj445670/c

LOCUS
DEFINITION Bj445670 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone ddv60b11 3', mRNA sequence.

ACCESSION
Bj445670

VERSION
Bj445670.1 GI:19420391

KEYWORDS
EST.

SOURCE
Dictyostelium discoideum

ORGANISM
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

REFERENCE
1 (bases 1 to 718)
Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
Full length cDNA of Dictyostelium discoideum at the vegetative
stage

JOURNAL
Unpublished (2002)

COMMENT
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855

FEATURES
source
Email: tshini@genes.nig.ac.jp.
location/Qualifiers
1. 718
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv60b11"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Alignment Scores:
Pred. No.: 0.118 Length: 718
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x Bj445670 (1-718)

Qy 1 HisAenHisAenHisAenHisAenHisAenHisAen 12
|||||

Db 158 CACATCATCATCATCATCATCATCATCATCAT 123

RESULT 83
Bj429204/c

LOCUS
DEFINITION Bj429204 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone ddv2013 3', mRNA sequence.

ACCESSION
Bj429204

VERSION
Bj429204.1 GI:19403926

KEYWORDS
EST.

SOURCE
Dictyostelium discoideum

ORGANISM
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

REFERENCE
1 (bases 1 to 723)
Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
Full length cDNA of Dictyostelium discoideum at the vegetative
stage

JOURNAL
Unpublished (2002)

COMMENT
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1. 723
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv2013"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Alignment Scores:
Pred. No.: 0.119 Length: 723
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x Bj429204 (1-723)

Qy 1 HisAenHisAenHisAenHisAenHisAenHisAen 12
|||||

Db 200 CACAAATCAATCAATCAATCAATCAATCAAT 165

RESULT 84
LOCUS BU434636/c 724 bp mRNA linear EST 13-MAR-2002
DEFINITION BU434636 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone dtdv24b15 3', mRNA sequence.
ACCESSION BU434636
VERSION BU434636.1 GI:19409358
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
1 (bases 1 to 724)
AUTHORS Unushihara, H., Tanaka, Y., Kohara, Y., and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..724
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dtdv24b15"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1lb="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Alignment Scores:
Pred. No.: 0.119 Length: 724
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BU434636 (1-724)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 202 CACAAATCAATCAATCAATCAATCAATCAAT 167

RESULT 85
LOCUS BU429573 725 bp mRNA linear EST 13-MAR-2002
DEFINITION BU429573 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone dtdv4c02 3', mRNA sequence.
ACCESSION BU429573
VERSION BU429573.1 GI:19404295
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
1 (bases 1 to 725)
AUTHORS Unushihara, H., Tanaka, Y., Kohara, Y., and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856

Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..725
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dtdv4c02"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1lb="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Alignment Scores:
Pred. No.: 0.119 Length: 725
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BU429573 (1-725)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 203 CACAAATCAATCAATCAATCAATCAATCAAT 168

RESULT 86
LOCUS CNB34352 725 bp mRNA linear EST 02-JUN-2004
DEFINITION AGENCOURT 24954489 NIH MGC 169 Mus musculus cDNA clone
IMAGE:30917704 5', mRNA sequence.
ACCESSION CNB34352
VERSION CNB34352.1 GI:47938924
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 725)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE NIH-MGC Consortium
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contract: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Jonathan Kuo, NIMH
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: NDCM384 row: f column: 17
High quality sequence stop: 619.

FEATURES
source
1..725
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30917704"
/lab_host="PH10B (TI-phage-resistant)"
/clone_1lb="NIH_MGC_169"
/note="Organ: Testicles; Vector: pDNR-LIB; Site 1: SfiI
(ggccataggcc); Site 2: SfiI (ggccggccggcc); cDNA made
by oligo-dT priming and directionally cloned. 5' and 3'
adaptors were used in cloning as follows:
5'-AACGAGTGTATTCACGACGAGTGGGCAATTACGGCGG-3' and
5'-ATTCTAGAGCGCCGAGCGGCCGACATG-CT(30)NN-3'. Full-length
enriched library was constructed using the Clontech

Creator SMART kit and size-selected to contain the 0.5 kb size fraction. Library created in the laboratory of M. Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."

ACCESSION CR277961
VERSION CR277961.1 GI:50056938
KEYWORDS GSS; genome survey sequence; MLCR,
SOURCE *Mus musculus* (house mouse)

KEYWORDS GSS; genome survey sequence; MICR. **SOURCE** Mus musculus (house mouse)

KEYWORDS GSS; genome survey sequence; MICR.
SOURCE Mus musculus (house mouse)

KEYWORDS GSS; genome survey sequence; MICR. **SOURCE** Mus musculus (house mouse)

BU401045/c 741 bp mRNA linear EST 10-MAR-2002
LOCUS BU401045 Dictyostelium discoideum cDNA library, SF Dictyostelium
DEFINITION dictyostelium cDNA clone dds15024 3', mRNA sequence.
ACCESSION BU401045
VERSION BU401045.1 GI:19313962
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.
REFERENCE 1 (bases 1 to 741)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE Full length cDNA of Dictyostelium discoideum at the slug stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@gene.nig.ac.jp.
Location/Qualifiers
1..741
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dds15024"
/sex="mat A"
/dev_stage="Slug stage"
/clone_lib="Dictyostelium discoideum cDNA library, SF"

ORIGIN
Alignment Scores:
Pred. No.: 0.122 Length: 741
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BU401045 (1-741)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
|||||
173 CACATCACATCACATCACATCACATCACAT 138

FEATURES
source
RESULT 93 749 bp DNA linear GSS 29-JAN-2003
BX211625
LOCUS BX211625 Dario rerio genomic clone DKEY-250M24, genomic survey sequence.
DEFINITION BX211625
ACCESSION BX211625.1 GI:28043511
VERSION GSS.
KEYWORDS Dario rerio (zebrafish)
SOURCE Dario rerio
ORGANISM Dario rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 749)
AUTHORS Humphray,S.J., Huckle,E. and Durham,J.L.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humphray@sanger.ac.uk unpublished
This sequence was generated from the T7 end of BAC 250M24. 250M24
is part of the Daniokey BAC Library created by R. Plastek and N.V.
Keygene. Further details:
http://www.sanger.ac.uk/Projects/D_rerio/.
Location/Qualifiers
1..749
/organism="Dario rerio"
/mol_type="genomic DNA"

FEATURES
source

/db_xref="taxon:7955"
/clone="DKEY-250M24"
/issue_type="Testis"
/note="Vector pIndigoBAC-536"

ORIGIN
Alignment Scores:
Pred. No.: 0.123 Length: 749
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-858-332g-15 (1-12) x BX211625 (1-749)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
|||||
485 CACATCACATCACATCACATCACATCACAT 450

FEATURES
source
RESULT 94 759 bp mRNA linear EST 13-MAR-2002
BU434227/c
LOCUS BU434227 Dictyostelium discoideum cDNA library, VF Dictyostelium
DEFINITION BU434227 dictyostelium cDNA clone dvl6c07 3', mRNA sequence.
ACCESSION BU434227
VERSION BU434227.1 GI:19408949
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.
REFERENCE 1 (bases 1 to 759)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@gene.nig.ac.jp.
Location/Qualifiers
1..759
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dvl6c07"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN
Alignment Scores:
Pred. No.: 0.125 Length: 759
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x BU434227 (1-759)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
|||||
208 CACATCACATCACATCACATCACATCACAT 173

FEATURES
source
RESULT 95 760 bp DNA linear GSS 05-JUN-2004
AG542244
LOCUS AG542244 Mus musculus molossinus DNA, clone:MM901-45M14.T7, genomic survey
DEFINITION Mus musculus molossinus DNA, clone:MM901-45M14.T7, genomic survey

sequence.
 AG542244 GI:48302658
 AG542244.1
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus molossinus
 Mus musculus molossinus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
 BAC end Sequences of Library MSNg01
 2 (bases 1 to 760)
 Unpublished
 Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
 Direct Submission
 Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the mouse BAC library MSNg01. For BAC
 library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
 Tsukuba Institute, Bio Resource Center,
 The Institute of Physical and Chemical Research (RIKEN) 3-1-1
 Koyadai, Tsukuba, 305-0074 Japan
 phone: 81-298-36-9189, fax: 81-298-36-9199
 e-mail: abe@rtc.riken.jp
 PRIMERS
 Sequencing : T7
 LIBRARY
 Vector : pBACe3.6
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI.

FEATURES
 source
 1..760
 location/Qualifiers
 /organism="Mus musculus molossinus"
 /mol_type="genomic DNA"
 /sub_species="molossinus"
 /db_xref="taxon:57486"
 /clone="MSNg01-455M14.T7"
 /sex="male"
 /tissue_type="mixture of kidney and spleen"
 /clone_lib="MSNg01 Mouse Male BAC Library"

ORIGIN
 Alignment Scores:
 Pred. No.: 0.125 Length: 760
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-858-332G-15 (1-12) x AG542244 (1-760)

QY
 1 H18A8NH18A8NH18A8NH18A8NH18A8N 12
 |||||
 Db 123 CACACCAACCAACCAACCAACCAACCAAC 158

RESULT 96
 CO427845 761 bp mRNA linear EST 06-JUL-2004
 LOCUS UI-M-HW0-CTP-a-10-0-UI.r1 NIH_BMAP_HW0 Mus musculus cDNA clone
 DEFINITION IMAGE:30679209 5', mRNA sequence.
 ACCESSION CO427845
 VERSION CO427845.1 GI:49674139
 KEYWORDS
 SOURCE EST.
 Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 761)
 NIH-MGC http://mgc.nci.nih.gov/.

TITLE
 JOURNAL
 COMMENT
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Straubeberg, Ph.D.
 Email: cgabs-@email.nih.gov
 Tissue Procurement: Dr. James Lin University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/nousefl.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pyX-5.

FEATURES
 source
 1..761
 location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30679209"
 /tissue_type="whole eye"
 /dev_stage="newborn (1, 5, 15 days)"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_HW0"
 /note="Organ: Eye; Vector: pyX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction. Ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pyX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AATTAATTACG. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): "Gene Discovery in the
 Developing Mouse Nervous System", supported by National
 Institute of Mental Health (NIMH)."

ORIGIN
 Alignment Scores:
 Pred. No.: 0.125 Length: 761
 Score: 84.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 7 Gaps: 0

US-09-858-332G-15 (1-12) x CO427845 (1-761)

QY
 1 H18A8NH18A8NH18A8NH18A8NH18A8N 12
 |||||
 Db 462 CACACCAACCAACCAACCAACCAACCAAC 497

RESULT 97
 BJ430651/c 762 bp mRNA linear EST 13-MAR-2002
 LOCUS BJ430651
 DEFINITION BJ430651 Dictyostelium discoideum cDNA library, VF Dictyostelium
 discoidem cDNA clone ddv8k01 3', mRNA sequence.
 ACCESSION BJ430651
 VERSION BJ430651.1 GI:19405373
 KEYWORDS
 SOURCE EST.
 Dictyostelium discoidem
 ORGANISM Dictyostelium discoidem
 Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 1 (bases 1 to 762)
 Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
 Full length cDNA of Dictyostelium discoidem at the vegetative
 stage
 Unpublished (2002)
 Contact: Tadao Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..762
Location/Qualifiers

/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv8k01"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1lb="Dictyostelium discoideum cDNA library, VF"

ALIGNMENT SCORES:

Pred. No.: 0.125 Length: 762
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BJ430651 (1-762)

Qy 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 203 CACATCACAATCACAATCACAATCACAATCACAAT 168

RESULT 98

LOCUS BJ431956/c

DEFINITION BJ431956 Dictyostelium discoideum cDNA library, VF EST 13-MAR-2002
discoidium cDNA clone ddv17105 3', mRNA sequence.

ACCESSION BJ431956

VERSION BJ431956.1 GI:19406678

KEYWORDS EST.

SOURCE Dictyostelium discoideum

ORGANISM Dictyostelium discoideum

REFERENCE Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.

AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.

TITLE Full length cDNA of Dictyostelium discoideum at the vegetative

JOURNAL stage

COMMENT Contact: Tadao Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1..763

/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv17105"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1lb="Dictyostelium discoideum cDNA library, VF"

ORIGIN

ALIGNMENT SCORES:

Pred. No.: 0.126 Length: 763
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BJ431956 (1-763)

Qy 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 201 CACATCACAATCACAATCACAATCACAATCACAAT 166

RESULT 99

LOCUS BJ432448/c

DEFINITION BJ432448 Dictyostelium discoideum cDNA library, VF EST 13-MAR-2002
discoidium cDNA clone ddv18p15 3', mRNA sequence.

ACCESSION BJ432448

VERSION BJ432448.1 GI:19407170

KEYWORDS EST.

SOURCE Dictyostelium discoideum

ORGANISM Dictyostelium discoideum

REFERENCE Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.

AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.

TITLE Full length cDNA of Dictyostelium discoideum at the vegetative

JOURNAL stage

COMMENT Contact: Tadao Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1..768

/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv18p15"
/sex="mat A"
/dev_stage="Growth phase"
/clone_1lb="Dictyostelium discoideum cDNA library, VF"

ORIGIN

ALIGNMENT SCORES:

Pred. No.: 0.126 Length: 768
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x BJ432448 (1-768)

Qy 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 202 CACATCACAATCACAATCACAATCACAATCACAAT 167

RESULT 100

LOCUS BJ431981/c

DEFINITION BJ431981 Dictyostelium discoideum cDNA library, VF EST 13-MAR-2002
discoidium cDNA clone ddv17a10 3', mRNA sequence.

ACCESSION BJ431981

VERSION BJ431981.1 GI:19406703

KEYWORDS EST.

SOURCE Dictyostelium discoideum

ORGANISM Dictyostelium discoideum

REFERENCE Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.

AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.

TITLE Full length cDNA of Dictyostelium discoideum at the vegetative

JOURNAL stage

COMMENT Contact: Tadao Shin-i

Center For Genetic Resource Information

National Institute of Genetics
1111 Yatae, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifier

FEATURES

source
1. .774
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv17a10"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VP"

ORIGIN

Alignment Scores:

Pred. No.:	0.127	Length:	774
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-858-332g-15 (1-12) x BJ431981 (1-774)

Qy	1	HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn	12
Db	206	CACATCACAATCACAATCACAATCACAATCACAAT	171

Search completed: October 4, 2005, 13:20:06
Job time : 3126 secs

THIS PAGE BLANK (USPTO)

C 86 59 70.2 43676 3 US-09-356-952-12
87 59 70.2 55195 4 US-09-949-016-15854
88 59 70.2 57218 4 US-09-949-016-15512
89 59 70.2 72549 4 US-09-949-016-15477
90 59 70.2 134008 4 US-09-949-016-15841
91 59 70.2 152524 4 US-09-949-016-12683
92 59 70.2 152524 4 US-09-949-016-11194
93 59 70.2 580073 4 US-08-545-528D-1
94 58 69.0 292 1 US-08-097-829-7
95 58 69.0 292 1 US-08-577-403-7
96 58 69.0 435 4 US-09-248-796A-9887
97 58 69.0 601 4 US-09-949-016-165982
98 58 69.0 601 4 US-09-949-016-165983
99 58 69.0 1431 4 US-09-248-796A-4891
100 58 69.0 1644 4 US-09-248-796A-858

Sequence 12, Appl
Sequence 15854, A
Sequence 16512, A
Sequence 16477, A
Sequence 13841, A
Sequence 12683, A
Sequence 13194, A
Sequence 1, Appl
Sequence 7, Appl
Sequence 9887, Ap
Sequence 165982,
Sequence 165983,
Sequence 4891, Ap
Sequence 858, App

ALIGNMENTS

RESULT 1
US-10-125-594-3/C

/ Sequence 3, Application US/10125594
/ Patent No. 6740747
/ GENERAL INFORMATION:

/ APPLICANT: Kaushik, Azad
/ TITLE OF INVENTION: No. 6740747e1 Bovine VDJ Casette, Bf1H1, Suitable for Antigenic
/ FILE REFERENCE: 12837-4
/ CURRENT APPLICATION NUMBER: US/10/125, 594
/ CURRENT FILING DATE: 2002-04-19
/ PRIOR APPLICATION NUMBER: US 60/284, 899
/ PRIOR FILING DATE: 2001-04-20
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 3
/ LENGTH: 115
/ TYPE: DNA
/ ORGANISM: Bovine
US-10-125-594-3

Alignment Scores:

Pred. No.: 0.000562 Length: 115
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-10-125-594-3 (1-115)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 80 CATACCATTAACCATTAACCATTAACCATTAAC 45

RESULT 2
US-08-453-477-1

/ Sequence 1, Application US/08453477
/ Patent No. 5529912
/ GENERAL INFORMATION:

/ APPLICANT: Henry, Susan A.
/ TITLE OF INVENTION: Inositol-Excreting Yeast
/ NUMBER OF SEQUENCES: 2
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Carnegie Mellon University
/ STREET: 4400 Forbes Avenue
/ CITY: Pittsburgh
/ STATE: Pennsylvania
/ COUNTRY: USA
/ ZIP: 15213
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 5-1/4 low density diskette
/ COMPUTER: IBM PC or compatibles

OPERATING SYSTEM: MS-DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,477

FILING DATE: 30-MAY-95

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/672,355

FILING DATE: 20-MAR-91

ATTORNEY/AGENT INFORMATION:

NAME: Frederick H. Cohen; Mary-Elizabeth Buckles

REGISTRATION NUMBER: 28,061,31,907

REFERENCE NUMBER: 93-168

TELECOMMUNICATION INFORMATION:

TELEPHONE: 412/288-4164

TELEFAX: 412/288-3063

TELEX: 27871

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1897 base pairs

TYPE: nucleic acid

STANDARDNESS: double

TOPOLOGY: linear

MOLECULE TYPE:

DESCRIPTION: Genomic DNA with translated open-reading

HYPOHETICAL: No frame containing the putative Opi1 protein

ANTI-SENSE: No

ORIGINAL SOURCE:

ORGANISM: Saccharomyces cerevisiae

STRAIN: Generic laboratory strain

IMMEDIATE SOURCE:

LIBRARY: YCp19 Library

CLONE: p(SPO11)1

POSITION IN GENOME:

CHROMOSOME/SEGMENT: Chromosome VIII

MAP POSITION: Adjacent to SPO11

FEATURE:

NAME/KEY: Opi1 Gene

LOCATION: GenBank/EMBL Data Bank. Accession

IDENTIFICATION METHOD: Opi1 gene was cloned using the

IDENTIFICATION METHOD: SPO11 mapping technique. The cloned gene complemented an

OTHER INFORMATION: The cloned Opi1 gene restored INO1

PUBLICATION INFORMATION: of overproduction of inositol.

AUTHORS: Michael J. White

AUTHORS: Jeanne P. Hirsch

AUTHORS: Susan A. Henry

TITLE: The Opi1 Gene of Saccharomyces cerevisiae, a

TITLE: Negative Regulator of Phospholipid Biosynthesis,

TITLE: Encodes a Protein Containing Polyglutamine

JOURNAL: The Journal of Biological Chemistry

VOLUME: 266

ISSUE: 2

PAGES: 863-872

DATE: 15-1-91

US-08-453-477-1

/

/

/

/

/

/

/

/

/

/

/

/

/

/

Alignment Scores:

Pred. No.: 0.0105 Length: 1897
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-858-332G-15 (1-12) x US-08-453-477-1 (1-1897)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 1730 CATATCATATCATATCATATCATATCATATCATAT 1765

RESULT 3

US-08-453-461-1

Sequence 1, Application US/08453461

Patent No. 5599701

GENERAL INFORMATION:

APPLICANT: Henry, Susan A.

APPLICANT: White, Michael J.

TITLE OF INVENTION: Inositol-Excreting Yeast

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Carnegie Mellon University

STREET: 4400 Forbes Avenue

CITY: Pittsburgh

STATE: Pennsylvania

COUNTRY: USA

ZIP: 15213

COMPUTER READABLE FORM:

MEDIUM TYPE: 5-1/4 low density diskette

COMPUTER: IBM PC or compatibles

OPERATING SYSTEM: MS-DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/453,461

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/165,182

FILING DATE: 02-DEC-93

APPLICATION NUMBER: US07640495

FILING DATE: 14-JAN-91

ATTORNEY/AGENT INFORMATION:

NAME: Frederick H. Cohen, Mary-Elizabeth Buckles

REGISTRATION NUMBER: 28,061; 31,907

REFERENCE/DOCKET NUMBER: 93-167

TELECOMMUNICATION INFORMATION:

TELEPHONE: 412/288-4164

TELEFAX: 412/288-3063

TELEX: 277871

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1897 base.pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE:

DESCRIPTION: Genomic DNA with translated open-reading

DESCRIPTION: frame containing the putative Opi1 protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: *Saccharomyces cerevisiae*

STRAIN: Generic laboratory strain

IMMEDIATE SOURCE:

LIBRARY: YC019 library

CLONE: p(SPO11)1

POSITION IN GENOME:

CHROMOSOME/SEGMENT: Chromosome VIII

MAP POSITION: Adjacent to SPO11

FEATURE:

NAME/KEY: Opi1 gene

LOCATION: GenBank/EMBL Data Bank. Accession

LOCATION: number J05727

IDENTIFICATION METHOD: Opi1 gene was cloned using the

IDENTIFICATION METHOD: SPO11 mapping technique. The cloned gene complemented an

OTHER INFORMATION: The cloned Opi1 gene restored INO1

OTHER INFORMATION: regulation in an opi1 mutant resulting in the loss

PUBLICATION INFORMATION:

AUTHORS: Michael J. White

AUTHORS: Jeanne P. Hirsch

AUTHORS: Susan A. Henry

TITLE: The Opi1 Gene of *Saccharomyces cerevisiae*, a

TITLE: Negative Regulator of Phospholipid Biosynthesis,

TITLE: Encodes a Protein Containing Polyglutamine

TITLE: Tracts and a Leucine Zipper

JOURNAL: The Journal of Biological Chemistry

VOLUME: 266

ISSUE: 2

PAGES: 863-872

DATE: 15-1-91

US-08-453-461-1

Alignment Scores:

Pred. No.: 0.0105 Length: 1897

Score: 84.00 Matches: 12

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 1 Gaps: 0

US-09-858-332g-15 (1-12) x US-08-453-461-1 (1-1897)

OY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 1730 CATATCATATCATATCATATCATATCATATCATAT 1765

RESULT 4

US-10-125-594-5/C

Sequence 5, Application US/10125594

Patent No. 6740747

GENERAL INFORMATION:

APPLICANT: Kaushik, Azad

APPLICANT: Saini, Surinder Singh

TITLE OF INVENTION: No. 6740747el Bovine VDJ Cassette, BFIHL, Suitable for Antigeniz

FILE REFERENCE: 12837-4

CURRENT APPLICATION NUMBER: US/10/125,594

CURRENT FILING DATE: 2002-04-19

PRIOR APPLICATION NUMBER: US 60/284,899

PRIOR FILING DATE: 2001-04-20

NUMBER OF SEQ ID NOS: 11

SOFTWARE: PatentIn version 3.1

SEQ ID NO 5

LENGTH: 139

TYPE: DNA

ORGANISM: Bovine

US-10-125-594-5

Alignment Scores:

Pred. No.: 0.00413 Length: 139

Score: 78.00 Matches: 11

Percent Similarity: 100.00% Conservative: 1

Best Local Similarity: 91.67% Mismatches: 0

Query Match: 92.86% Indels: 0

DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x US-10-125-594-5 (1-139)

OY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 56 TACAACCATTAACCATTAACCATTAACCATTAAC 21

RESULT 5

US-10-125-594-2/C

Sequence 2, Application US/10125594

Patent No. 6740747

GENERAL INFORMATION:

APPLICANT: Kaushik, Azad

APPLICANT: Saini, Surinder Singh

TITLE OF INVENTION: No. 6740747el Bovine VDJ Cassette, BFIHL, Suitable for Antigeniz

FILE REFERENCE: 12837-4

CURRENT APPLICATION NUMBER: US/10/125,594

CURRENT FILING DATE: 2002-04-19

QY 1 HisAenHisAenHisAenHisAenHis 11
DB 2487 CATATCATATCATATCATATCATATCAT 2519

RESULT 12
US-09-323-433A-3
; Sequence 3, Application US/09323433A
; Patent No. 6218512
; GENERAL INFORMATION:
; APPLICANT: Guarente, Leonard P.
; APPLICANT: Austriaco Jr., Nicanor
; APPLICANT: Claus, James J.
; APPLICANT: Cole, Francesca
; APPLICANT: Kennedy, Brian
; TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESENCE IN
; FILE REFERENCE: 0050.1491-003
; CURRENT APPLICATION NUMBER: US/09/323,433A
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 08/396,001
; PRIOR FILING DATE: 1995-02-28
; PRIOR APPLICATION NUMBER: PCT/US94/09351
; PRIOR FILING DATE: 1994-08-15
; PRIOR APPLICATION NUMBER: US 08/107,408
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 3455
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (663)...(3164)
; OTHER INFORMATION: UTH4
US-09-323-433A-3

Alignment Scores:
Pred. No.: 0.118 Length: 3455
Score: 78.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.86% Indels: 0
DB: 3 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-323-433A-3 (1-3455)

QY 1 HisAenHisAenHisAenHisAenHis 11
DB 2487 CATATCATATCATATCATATCATATCAT 2519

RESULT 13
US-09-826-752-3
; Sequence 3, Application US/09826752
; Patent No. 6787300
; GENERAL INFORMATION:
; APPLICANT: Guarente, Leonard P.
; APPLICANT: Austriaco Jr., Nicanor
; APPLICANT: Claus, James J.
; APPLICANT: Cole, Francesca
; APPLICANT: Kennedy, Brian
; TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESENCE IN
; FILE REFERENCE: 0050.1491-005
; CURRENT APPLICATION NUMBER: US/09/826,752
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 08/396,001
; PRIOR FILING DATE: 1995-02-28
; PRIOR APPLICATION NUMBER: PCT/US94/09351
; PRIOR FILING DATE: 1994-08-15
; PRIOR APPLICATION NUMBER: US 08/107,408
; PRIOR FILING DATE: 1993-08-16

PRIOR APPLICATION NUMBER: US 09/323,433
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3455
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (663)...(3164)
; OTHER INFORMATION: UTH4
US-09-826-752-3

Alignment Scores:
Pred. No.: 0.118 Length: 3455
Score: 78.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.86% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-826-752-3 (1-3455)

QY 1 HisAenHisAenHisAenHisAenHis 11
DB 2487 CATATCATATCATATCATATCATATCAT 2519

RESULT 14
US-09-248-796A-7109
; Sequence 7109, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 7109
; LENGTH: 921
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-7109

Alignment Scores:
Pred. No.: 0.0401 Length: 921
Score: 77.00 Matches: 11
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 91.67% Mismatches: 0
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-7109 (1-921)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
DB 478 CATATCATATCAACCAATATCAACCAACCA 513

RESULT 15
US-09-792-024-7
; Sequence 7, Application US/09792024
; Patent No. 6783985
; GENERAL INFORMATION:
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard

```

; TITLE OF INVENTION: Gene Disruption Methodologies for Drug
; TITLE OF INVENTION: Targets Discovery
; FILE REFERENCE: 10182-004-999
; CURRENT APPLICATION NUMBER: US/09/792,024
; PRIOR FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 490
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1521
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-792-024-7

Alignment Scores:
Pred. No.: 0.0914 Length: 1521
Score: 76.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 90.48% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x US-09-792-024-7 (1-1521)

QY 2 AsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 173 AACCAACCAACCAACCAACCAACCAACCAAC 205

RESULT 16
US-09-248-796A-184
; Sequence 184, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 184
; LENGTH: 1806
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-184

Alignment Scores:
Pred. No.: 0.109 Length: 1806
Score: 76.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 90.48% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x US-09-248-796A-184 (1-1806)

QY 2 AsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 1147 AACCATACCAACCAACCAACCAACCAACCAAC 1179

RESULT 17
US-09-794-384A-3/C
; Sequence 3, Application US/09794384A
; Patent No. 6743968
; GENERAL INFORMATION:
; APPLICANT: Dellaporta, Stephen L.
; APPLICANT: Moreno, Maria A.
; APPLICANT: Yale University
; TITLE OF INVENTION: Methods and Compositions to Reduce or Eliminate
; TITLE OF INVENTION: Transmission of a Transgene
```

```

; FILE REFERENCE: 44574-5078-US
; CURRENT APPLICATION NUMBER: US/09/794,384A
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,524
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 4565
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Transposable element Ac
US-09-794-384A-3

Alignment Scores:
Pred. No.: 0.288 Length: 4565
Score: 76.00 Matches: 11
Percent Similarity: 91.67% Conservative: 0
Best Local Similarity: 91.67% Mismatches: 1
Query Match: 90.48% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x US-09-794-384A-3 (1-4565)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 3146 CACACCAACCAACCAACCAACCAACCAACCAAC 3111

RESULT 18
US-09-248-796A-1902
; Sequence 1902, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1902
; LENGTH: 747
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-1902

Alignment Scores:
Pred. No.: 0.0792 Length: 747
Score: 74.00 Matches: 10
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 88.10% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x US-09-248-796A-1902 (1-747)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 91 CATATACCAACCAACCAATCATCATCATCATCAT 126

RESULT 19
US-09-248-796A-4640
; Sequence 4640, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; APPLICANT: Yale University
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
```

```

; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4640
; LENGTH: 1953
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-4640

Alignment Scores:
Pred. No.:
Score: 0.393 Length: 1953
Percent Similarity: 72.00 Matches: 10
Best Local Similarity: 100.00% Conservative: 2
Query Match: 83.33% Mismatches: 0
DB: 85.71% Indels: 0
Gaps: 0

US-09-858-332g-15 (1-12) x US-09-248-796A-4640 (1-1953)
QY 1 HisAnHiSaEnHiSaEnHiSaEnHiSaEn 12
Db 1685 AACATCACAATCACAATCACAATCACAATCACAAGT 1720

RESULT 20
US-09-248-796A-4342
; Sequence 4342, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4342
; LENGTH: 903
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-4342

Alignment Scores:
Pred. No.:
Score: 0.237 Length: 903
Percent Similarity: 71.00 Matches: 10
Best Local Similarity: 100.00% Conservative: 1
Query Match: 90.91% Mismatches: 0
DB: 84.52% Indels: 0
Gaps: 0

US-09-858-332g-15 (1-12) x US-09-248-796A-4342 (1-903)
QY 2 AsnHiSaEnHiSaEnHiSaEnHiSaEnHiSaEn 12
Db 305 AACCAACCAACCAACCAACCAACCAACCAAC 337

RESULT 21
US-09-248-796A-8980
; Sequence 8980, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
```

```

; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 8980
; LENGTH: 192
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-8980

Alignment Scores:
Pred. No.:
Score: 0.0636 Length: 192
Percent Similarity: 70.00 Matches: 10
Best Local Similarity: 100.00% Conservative: 2
Query Match: 83.33% Mismatches: 0
DB: 83.33% Indels: 0
Gaps: 0

US-09-858-332g-15 (1-12) x US-09-248-796A-8980 (1-192)
QY 1 HisAnHiSaEnHiSaEnHiSaEnHiSaEn 12
Db 114 AACCAACCAACCAACCAACCAACCAACCAATAC 149

RESULT 22
US-09-248-796A-1110
; Sequence 1110, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBIC
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1110
; LENGTH: 966
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-1110

Alignment Scores:
Pred. No.:
Score: 0.343 Length: 966
Percent Similarity: 70.00 Matches: 10
Best Local Similarity: 100.00% Conservative: 1
Query Match: 90.91% Mismatches: 0
DB: 83.33% Indels: 0
Gaps: 0

US-09-858-332g-15 (1-12) x US-09-248-796A-1110 (1-966)
QY 2 AsnHiSaEnHiSaEnHiSaEnHiSaEnHiSaEn 12
Db 650 AATTACATCACAATCACAATCACAATCACAAT 682

RESULT 23
US-09-248-796A-12604
; Sequence 12604, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
```

```

US-09-858-332G-15 (1-12) x US-09-248-796A-12604 (1-1995)
Db 1295 AACCCACGACGACCAACCAACCAACCAACCAAC 1327
RESULT 24
US-09-817-310-1/c
Sequence 1, Application US/09817310
Patent No. 6534311
GENERAL INFORMATION:
APPLICANT: Stewart, Mary
APPLICANT: Kozma, Sarah
APPLICANT: Thomas, George
TITLE OF INVENTION: Drosophila Melanogaster P70 S6 Kinase
FILE REFERENCE: 4-20971/A
CURRENT APPLICATION NUMBER: US/09/817,310
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: 09/230,247
PRIOR FILING DATE: 1999-04-16
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 2556
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-817-310-1
Alignment Scores:
Pred. No.: 0.948 Length: 2556
Score: 70.00 Matches: 10
Percent Similarity: 91.61% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 83.33% Indels: 0
Gaps: 0
US-09-858-332G-15 (1-12) x US-09-817-310-1 (1-2556)
Qy 1 HisanHisanHisanHisanHisanHisanHisan 12
Db 2325 CATATCATATCATATCATATCATATCATATTC 2290
RESULT 25
US-10-355-724A-1/c
Sequence 1, Application US/10355724A
Patent No. 6833255
GENERAL INFORMATION:
APPLICANT: Stewart, Mary
APPLICANT: Kozma, Sarah
APPLICANT: Thomas, George
TITLE OF INVENTION: Drosophila Melanogaster P70 S6 Kinase
FILE REFERENCE: 4-20971/A
CURRENT APPLICATION NUMBER: US/10/355,724A

```

```

CURRENT FILING DATE: 2003-01-31
PRIOR APPLICATION NUMBER: GB 96/5498.4
PRIOR FILING DATE: 1996-07-24
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 2556
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-10-355-724A-1

Alignment Scores:
Pred. No.: 0.948 Length: 2556
Score: 70.00 Matches: 10
Percent Similarity: 91.67% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 83.33% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-10-355-724A-1 (1-2556)

Cy 1 Hisaenhisaenhisaenhisaenhisaenhisaen 12
|||||
Db 2325 CATATCATATCATATCATATCATATCATATCATAT 2290
|||||

RESULT 26
US-08-928-361B-2
Sequence 2, Application US/08928361B
Patent No. 6071518
GENERAL INFORMATION:
APPLICANT: Petersen, Carolyn
TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
TITLE OF INVENTION: SPECIES INFECTIONS
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETERS, VERNY, JONES & BIKSA
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-1840
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,361B
FILING DATE: 12-SEP-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Verny, Hana
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480.76-1(HV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-1677
TELEFAX: 650-324-1678
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-928-361B-2

Alignment Scores:
Pred. No.: 2.11 Length: 5511

```

Score: 70.00 Matches: 10
Percent Similarity: 91.67% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 83.33% Indels: 0
DB: 3 Gaps: 0

US-09-858-332g-15 (1-12) x US-08-928-361b-2 (1-5511)

QY 1 HisAenHisAenHisAenHisAenHisAenHisAen 12
DB 1443 TACAACCAACCAACCAACCAACCAACCAAC 1478

RESULT 27

US-09-588-995A-2
Sequence 2, Application US/09588995A
Patent No. 6514697

GENERAL INFORMATION:
APPLICANT: PETERSEN, CAROLYN
APPLICANT: BARNES, DEBRA A.
APPLICANT: NELSON, RICHARD C.
APPLICANT: GUT, JIRI

TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND
TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM
FILE REFERENCE: 480.19-5

CURRENT APPLICATION NUMBER: US/09/588, 995A

PRIOR FILING DATE: 2000-06-06

PRIOR FILING DATE: 1997-03-27

PRIOR FILING DATE: 1997-09-12

PRIOR FILING DATE: 1996-08-14

PRIOR FILING DATE: 1995-04-03

NUMBER OF SEQ ID NOS: 115

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 2

LENGTH: 5511

TYPE: DNA

ORGANISM: Cryptosporidium parvum

US-09-588-995A-2

Alignment Scores:

Pred. No.: 2.11 Length: 5511

Score: 70.00 Matches: 10

Percent Similarity: 91.67% Conservative: 1

Best Local Similarity: 83.33% Mismatches: 1

Query Match: 83.33% Indels: 0

DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x US-09-588-995A-2 (1-5511)

QY 1 HisAenHisAenHisAenHisAenHisAenHisAen 12

DB 1443 TACAACCAACCAACCAACCAACCAACCAAC 1478

RESULT 28

US-08-928-361b-1
Sequence 1, Application US/08928361B
Patent No. 6071518

GENERAL INFORMATION:
APPLICANT: Petersen, Carolyn
APPLICANT: PETERSEN, CAROLYN
APPLICANT: BARNES, DEBRA A.
APPLICANT: NELSON, RICHARD C.
APPLICANT: GUT, JIRI

TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:
ADDRESS: PETERS, VERNY, JONES & BIRSA
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: CA

COUNTRY: USA
ZIP: 94306-1840
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,361B
FILING DATE: 12-SEP-1997

CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/026,062

FILING DATE: 13-SEP-1996

ATTORNEY/AGENT INFORMATION:
NAME: Verry, Hana

REGISTRATION NUMBER: 30,518

REFERENCE/DOCKET NUMBER: 480.76-1(HV)

TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-1677

TELEFAX: 650-324-1678

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 7334 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-928-361b-1

Alignment Scores:

Pred. No.: 2.85 Length: 7334

Score: 70.00 Matches: 10

Percent Similarity: 91.67% Conservative: 1

Best Local Similarity: 83.33% Mismatches: 1

Query Match: 83.33% Indels: 0

DB: 3 Gaps: 0

US-09-858-332g-15 (1-12) x US-08-928-361b-1 (1-7334)

QY 1 HisAenHisAenHisAenHisAenHisAenHisAen 12

DB 3111 TACAACCAACCAACCAACCAACCAACCAAC 3146

RESULT 29

US-09-588-995A-1
Sequence 1, Application US/09588995A
Patent No. 6514697

GENERAL INFORMATION:
APPLICANT: PETERSEN, CAROLYN
APPLICANT: BARNES, DEBRA A.
APPLICANT: NELSON, RICHARD C.
APPLICANT: GUT, JIRI

TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND
TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM
FILE REFERENCE: 480.19-5

CURRENT APPLICATION NUMBER: US/09/588, 995A

PRIOR FILING DATE: 2000-06-06

PRIOR FILING DATE: 1997-03-27

PRIOR FILING DATE: 1997-09-12

PRIOR FILING DATE: 1996-08-14

PRIOR FILING DATE: 1995-04-03

NUMBER OF SEQ ID NOS: 115

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 1

LENGTH: 7334

TYPE: DNA

ORGANISM: Cryptosporidium parvum

```
US-09-588-995A-1
Alignment Scores:
Pred. No.: 2.85 Length: 7334
Score: 70.00 Matches: 10
Percent Similarity: 91.67% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 83.33% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-588-995A-1 (1-7334)
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 3111 TACACCAACACCAACCAACCAACCAACCAACCAAC 3146

RESULT 30
US-09-949-016-15086
; Sequence 15086, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15086
; LENGTH: 219964
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15086

Alignment Scores:
Pred. No.: 99.1 Length: 219964
Score: 70.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 83.33% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-949-016-15086 (1-219964)
QY 2 AsnHisAsnHisAsnHisAsnHisAsnHis 11
Db 181126 AATCATATCATATCATATCATATCATATCAT 181155

RESULT 31
US-09-949-016-16923
; Sequence 16923, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: US/09/949,016
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16923
; LENGTH: 390416
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16923

Alignment Scores:
Pred. No.: 180 Length: 390416
Score: 70.00 Matches: 10
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 83.33% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-949-016-16923 (1-390416)
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 205719 AATATCATATCATATCATATCATATATATATAT 205754

RESULT 32
US-09-248-796A-5488
; Sequence 5488, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 5488
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-5488

Alignment Scores:
Pred. No.: 0.827 Length: 1263
Score: 68.00 Matches: 9
Percent Similarity: 91.67% Conservative: 2
Best Local Similarity: 75.00% Mismatches: 1
Query Match: 80.95% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-5488 (1-1263)
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 377 CACAGCCACACACACACACACACACACACAGCATCAT 412

RESULT 33
US-09-248-796A-1205
; Sequence 1205, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT FILING DATE: US/09/248,796A
; PRIOR APPLICATION NUMBER: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
```

```
; SEQ ID NO 1205
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-1205

Alignment Scores:
Pred. No.: 0.596      Length: 600
Score: 66.50         Matches: 11
Percent Similarity: 70.59% Conservative: 1
Best Local Similarity: 64.71% Mismatches: 0
Query Match: 79.17% Indels: 5
DB: 4 Gaps: 1

US-09-858-332G-15 (1-12) x US-09-248-796A-1205 (1-600)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
   |||:|||||:|||||:|||||:|||||:|||||:
Db 161 CACAGTCACATCACAATCTGAGAACAAATCACAATCAATCAAT 211

RESULT 34
US-09-248-796A-8010
; Sequence 8010, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 8010
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (346), (379)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unkno
US-09-248-796A-8010

Alignment Scores:
Pred. No.: 0.563      Length: 492
Score: 66.00         Matches: 9
Percent Similarity: 91.67% Conservative: 2
Best Local Similarity: 75.00% Mismatches: 1
Query Match: 78.57% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-8010 (1-492)

QY 1 HisAenHisAenHisAenHisAenHisAen 12
   |||:|||||:|||||:|||||:|||||:|||||:
Db 154 CATCACATCACCATCATCATCTCTCATATCATATAT 189

RESULT 35
US-09-248-796A-333
; Sequence 333, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
```

```
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 333
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-333

Alignment Scores:
Pred. No.: 1.26      Length: 1065
Score: 66.00         Matches: 9
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 81.82% Mismatches: 0
Query Match: 78.57% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-333 (1-1065)

QY 2 AenHisAenHisAenHisAenHisAenHisAen 12
   ::|||:|||||:|||||:|||||:|||||:
Db 593 AGTCACACACACACACACACACACACACACACAC 625

RESULT 36
US-09-248-796A-6847
; Sequence 6847, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 6847
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-6847

Alignment Scores:
Pred. No.: 1.61      Length: 1344
Score: 66.00         Matches: 9
Percent Similarity: 91.67% Conservative: 2
Best Local Similarity: 75.00% Mismatches: 1
Query Match: 78.57% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-6847 (1-1344)

QY 1 HisAenHisAenHisAenHisAenHisAenHisAen 12
   |||:|||||:|||||:|||||:|||||:|||||:
Db 181 CATCACATCACCATCATCATCTCTCATATCATATAT 216

RESULT 37
US-08-700-651-1
; Sequence 1, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: LEECH, JAMES
; APPLICANT: NELSON, RICHARD, C.
; APPLICANT: GUY, JIRI
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAs AND RNAs
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
; FILE REFERENCE: 480.19-4 (HV)
; CURRENT APPLICATION NUMBER: US/08/700,651B
```



```

CURRENT FILING DATE: 1997-08-14
EARLIER APPLICATION NUMBER: 08/415,751
EARLIER FILING DATE: 1995-04-03
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 5163
TYPE: DNA
ORGANISM: Cryptosporidium parvum
US-08-700-651-1

Alignment Scores:
Pred. No.:      8.83      Length:      5163
Score:          65.00      Matches:      11
Percent Similarity: 66.67%  Conservative: 1
Best Local Similarity: 61.11%  Mismatches: 0
Query Match:      77.38%      Indels:      6
DB:               3          Gaps:         1

US-09-858-332G-15 (1-12) x US-08-700-651-1 (1-5163)

Qy      1 HisAsnHisAsnHis-----AsnHisAsnHisAsnHis 12
        |||||:|||||
Db      856 CACAACCTACCAACCCACACACTACCAACCCACACACACACACACAC 909
        |||||:|||||

RESULT 38
US-08-928-361B-4
: Sequence 4, Application US/08928361B
: Patent No. 6071518
: GENERAL INFORMATION:
:   APPLICANT: Petersen, Carolyn
:   TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
:   TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
:   TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
:   NUMBER OF SEQUENCES: 30
: CORRESPONDENCE ADDRESS:
:   ADDRESSEE: PETERS, VERNY, JONES & BIKSA
:   STREET: 385 Sherman Avenue, Suite 6
:   CITY: Palo Alto
:   STATE: CA
:   COUNTRY: USA
:   ZIP: 94306-1840
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/08/928,361B
:   FILING DATE: 12-SEP-1997
:   CLASSIFICATION:
:   PRIOR APPLICATION DATA:
:   APPLICATION NUMBER: US 60/026,062
:   FILING DATE: 13-SEP-1996
:   ATTORNEY/AGENT INFORMATION:
:   NAME: Verry, Hana
:   REGISTRATION NUMBER: 30,518
:   REFERENCE/DOCKET NUMBER: 480.76-1(HV)
:   TELECOMMUNICATION INFORMATION:
:   TELEPHONE: 650-324-1677
:   TELEFAX: 650-324-1678
:   INFORMATION FOR SEQ ID NO: 4:
:   SEQUENCE CHARACTERISTICS:
:   LENGTH: 5163 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: double
:   TOPOLOGY: linear
:   MOLECULE TYPE: DNA (genomic)
:   US-08-928-361B-4

Alignment Scores:
Pred. No.:      8.83      Length:      5163

```

```

Score: 65.00 Matches: 11
Percent Similarity: 66.67% Conservative: 0
Best Local Similarity: 61.1% Mismatches: 1
Query Match: 77.38% Indels: 6
DB: 3 Gaps: 1

US-09-858-332G-15 (1-12) x US-08-928-361B-4 (1-5163)
QY 1 HlaSbNHlaSbNHls-----AeNHlASbNHlASbNHlASbn 12
|||:|||:|||||
855 CACAACACTACAACCAACAACCTACAACCAACAACCAACAACCAACCAACCAAC 908

RESULT 39
US-09-588-995A-4
; Sequence 4, Application US/09588995A
; Patent No. 6514697
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: BARNES, DEBRA A.
; APPLICANT: NELSON, RICHARD C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND
; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: INFECTIONS
FILE REFERENCE: 480.19-5
CURRENT APPLICATION NUMBER: US/09/588,995A
PRIOR FILING DATE: 2000-06-06
PRIOR APPLICATION NUMBER: 08/827,171
PRIOR FILING DATE: 1997-03-27
PRIOR APPLICATION NUMBER: 08/928,361
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 08/700,651
PRIOR FILING DATE: 1996-08-14
PRIOR APPLICATION NUMBER: 08/415,751
PRIOR FILING DATE: 1995-04-03
NUMBER OF SEQ ID NOS: 115
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 5163
TYPE: DNA
ORGANISM: Cryptosporidium parvum
US-09-588-995A-4

Alignment Scores:
Pred. No.: 8.83 Length: 5163
Score: 65.00 Matches: 11
Percent Similarity: 66.67% Conservative: 1
Best Local Similarity: 61.1% Mismatches: 0
Query Match: 77.38% Indels: 6
DB: 4 Gaps: 1

US-09-858-332G-15 (1-12) x US-09-588-995A-4 (1-5163)
QY 1 HlaSaNHlaSbNHls-----AeNHlASbNHlASbNHlASbn 12
|||:|||:|||||
855 CACAACACTACAACCAACAACCTACAACCAACAACCAACAACCAACCAAC 908

RESULT 40
US-08-700-651-2
; Sequence 2, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: LEECH, JAMES
; APPLICANT: NELSON, RICHARD, C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAs AND RNAs
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
; FILE REFERENCE: 480.19-4(HV)
CURRENT APPLICATION NUMBER: US/08/700,651B
PRIOR FILING DATE: 1997-08-14
EARLIER APPLICATION NUMBER: 08/415,751

```



```
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-11186

Alignment Scores:
Pred. No.: 0.453 Length: 225
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x US-09-248-796A-11186 (1-225)

QY 1 Hisanhi$anhi$anhi$anhi$anhi$an 12
DB 9 CATACCATTAACCAACCAACCAACCAACCAAC 44

RESULT 44
US-09-248-796A-8699
/ Sequence 8699, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstock et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 8699
/ LENGTH: 465
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-8699

Alignment Scores:
Pred. No.: 0.966 Length: 465
Score: 64.00 Matches: 2
Percent Similarity: 91.67% Conservative: 9
Best Local Similarity: 75.00% Mismatches: 1
Query Match: 76.19% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x US-09-248-796A-8699 (1-465)

QY 1 Hisanhi$anhi$anhi$anhi$anhi$an 12
DB 193 CACACCAACAACCCCAATCAACCAACGATCAGT 228

RESULT 45
US-09-248-796A-4338
/ Sequence 4338, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstock et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 4338
/ LENGTH: 1254
/ TYPE: DNA
```

```
/ ORGANISM: Candida albicans
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (50), (51), (52), (53), (54), (55), (59), (63), (65), (71)
/ OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknown
US-09-248-796A-4338

Alignment Scores:
Pred. No.: 2.72 Length: 1254
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x US-09-248-796A-4338 (1-1254)

QY 1 Hisanhi$anhi$anhi$anhi$anhi$an 12
DB 73 CATATACCAACCAACCAACCAATCATCATCAATAT 108

RESULT 46
US-09-596-002-21
/ Sequence 21, Application US/09596002
/ Patent No. 6632636
/ GENERAL INFORMATION:
/ APPLICANT: Lagace, Robert, E.
/ APPLICANT: Patterson, Chandra
/ TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
/ FILE REFERENCE: PM-0008-4 US
/ CURRENT APPLICATION NUMBER: US/09/596,002
/ PRIOR FILING DATE: 2000-06-16
/ PRIOR APPLICATION NUMBER: 60/140,121
/ PRIOR FILING DATE: 1999-06-18
/ NUMBER OF SEQ ID NOS: 41
/ SOFTWARE: PERL Program
/ SEQ ID NO 21
/ LENGTH: 39003
/ TYPE: DNA
/ ORGANISM: M. catarrhalis
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte template ID No. 6632636 21
/ PUBLICATION INFORMATION:
US-09-596-002-21

Alignment Scores:
Pred. No.: 98.3 Length: 39003
Score: 64.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 4 Gaps: 0

US-09-858-332g-15 (1-12) x US-09-596-002-21 (1-39003)

QY 1 Hisanhi$anhi$anhi$anhi$anhi$ 9
DB 3052 CACAATCATATCATATCATATCAACCAT 3078

RESULT 47
US-09-949-016-12572
/ Sequence 12572, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: C1001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
```

```

? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO: 12572
? LENGTH: 53806
? TYPE: DNA
? ORGANISM: Human
?
US-03-949-016-12572

```

Alignment Scores:	
Pred. No.:	137
Score:	64.00%
Percent Similarity:	100.00%
Best Local Similarity:	66.67%
Query Match:	76.19%
DB:	4
	Indels: 0
	Gaps: 0
	Caps: 0
	Length: 53906
	Matches: 8
	Mismatches: 0
	Conservative: 4
	Misconservative: 0
	Indels: 0
	Gaps: 0
	Caps: 0

US-09-858-332G-15 (1-12) X US-09-949-016-12572 (1-53806)

QY 1 HisaSnHisasnHisasSnHisasnHisasn 12
 |||||:::|||||:::|||::|||::
 Db 22035 CACAATCACCACACACATCACCATCACCATCAC 22070

RESULT 48
 US-09-248-796A-12843/C
 ; Sequence 12843, Application US/09248796A
 ; Patent No. 6747113
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstein et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 12843
 ; LENGTH: 216
 ; TYPE: DNA
 ; ORGANISM: Candida albicans
 ; US-09-248-796A-12843

Alignment Scores:	
Pred. No.:	0.585
Score:	63.00
Percent Similarity:	83.33%
Best local Similarity:	75.00%
Query Match:	75.00%
DB:	4
Length:	216
Matches:	9
Conservative:	1
Mismatches:	2
Indels:	0
Gaps:	0

US-09-858-332G-15 (1-12) X US-09-248-796A-12843 (1-216)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
 |||||
 Db 160 CACAAATCAGAATCATATACACATCACCACCGGAC 125

RESULT 49
US-09-248-796A-12841
; Sequence 12841, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A

```

1  CURRENT FILING DATE: 1999-02-12
2  PRIOR APPLICATION NUMBER: US 60/074,722
3  PRIOR FILING DATE: 1998-02-13
4  PRIOR APPLICATION NUMBER: US 60/096,409
5  PRIOR FILING DATE: 1998-08-13
6  NUMBER OF SEQ ID NOS: 28208
7  SEQ ID NO 12841
8  LENGTH: 330
9  TYPE: DNA
10 ORGANISM: Candida albicans
11 US-09-248-796A-12841

```

Alignment Scores:	
Pred. No.:	0.911
Score:	63.00
Percent Similarity:	83.33%
Best Local Similarity:	75.00%
Query Match:	75.00%
DB:	4
Length:	330
Matches:	9
Conservative:	2
Mismatches:	2
Indels:	0
Gaps:	0

US-09-858-332G-15 (1-12) X US-09-248-796A-12841 (1-330)

[illegible]

```

1  RESULT 50
2  US-09-949-016-103758/C
3  ; Sequence 103758, Application US/09949016
4  ; Patent No. 6812339
5  ; GENERAL INFORMATION:
6  ; APPLICANT: VENTER, J. Craig et al.
7  ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
8  ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
9  ; FILE REFERENCE: CLO01007
10 ; CURRENT APPLICATION NUMBER: US/09/949,016
11 ; CURRENT FILING DATE: 2000-04-14
12 ; PRIOR APPLICATION NUMBER: 60/241,755
13 ; PRIOR FILING DATE: 2000-10-20
14 ; PRIOR APPLICATION NUMBER: 60/237,768
15 ; PRIOR FILING DATE: 2000-10-03
16 ; PRIOR APPLICATION NUMBER: 60/231,498
17 ; PRIOR FILING DATE: 2000-09-08
18 ; NUMBER OF SEQ ID NOS: 207012
19 ; SOFTWARE: FastSeq for Windows Version 4.0
20 ; SEQ ID NO 103758
21 ; LENGTH: 522

```

```

ORGANISM: Human
BUS-03-949-016-103758

Alignment Scores:
Pred. No.:
Score: 1.47
Length: 522
Percent Similarity: 63.00
Matches: 8
Conservative: 3
Mismatch: 1
Indels: 0
Query Match: 75.00%
Gaps: 0

```

US-09-858-332G-15 (1-12) X US-09-949-016-103758 (1-522)

1 Hisaenhiisaenhiisaenhiisaenhiisaenhiisaen 12
445 CAGACACACACACACATATACACAGCACACACCATCAC 410

RESULT 51
 US-09-248-796A-5078
 Sequence 5078, Application US/09248796A
 Patent No. 6747137
 GENERAL INFORMATION:
 APPLICANT: Keith Weinstein et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

```
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 5078
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Candida albicans
; US-09-248-796A-5078

Alignment Scores:
Pred. No.: 2,02      Length: 708
Score: 63.00      Matches: 9
Percent Similarity: 91.67%      Conservative: 2
Best Local Similarity: 75.00%      Mismatches: 1
Query Match: 75.00%      Indels: 0
DB: 4      Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-5078 (1-708)
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 425 CACACCACCAACCAACCAACCAACCAACCAACCAAC 460

RESULT 52
US-09-248-796A-4979
; Sequence 4979, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4979
; LENGTH: 2577
; TYPE: DNA
; ORGANISM: Candida albicans
; US-09-248-796A-4979

Alignment Scores:
Pred. No.: 7,78      Length: 2577
Score: 63.00      Matches: 8
Percent Similarity: 91.67%      Conservative: 3
Best Local Similarity: 66.67%      Mismatches: 1
Query Match: 75.00%      Indels: 0
DB: 4      Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-4979 (1-2577)
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 467 CACAGTCACATCATCATCATCATCATCATCATCATCAT 502

RESULT 53
US-09-949-016-14613/c
; Sequence 14613, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
```

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14613
; LENGTH: 30656
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(30656)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-14613

Alignment Scores:
Pred. No.: 103      Length: 30656
Score: 63.00      Matches: 8
Percent Similarity: 91.67%      Conservative: 3
Best Local Similarity: 66.67%      Mismatches: 1
Query Match: 75.00%      Indels: 0
DB: 4      Gaps: 0

US-09-858-332G-15 (1-12) x US-09-949-016-14613 (1-30656)
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 3770 CACACCACCAACCAACCAACCAACCAACCAACCAAC 3735

RESULT 54
US-09-949-016-11926
; Sequence 11926, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11926
; LENGTH: 95750
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(95750)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-11926

Alignment Scores:
Pred. No.: 338      Length: 95750
Score: 63.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 3
Best Local Similarity: 72.73%      Mismatches: 0
Query Match: 75.00%      Indels: 0
DB: 4      Gaps: 0

US-09-858-332G-15 (1-12) x US-09-949-016-11926 (1-95750)
```

```

QY 1 HisAenHisAenHisAenHisAenHis 11
Db 65017 CATATACACCATCATCATATCATCATCAT 65049

RESULT 55
US-09-663-600A-149
/ Sequence 149, Application US/09663600A
/ Patent No. 6573068
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, Jean-Baptiste
/ APPLICANT: Duclert, Aymeric
/ APPLICANT: Bouqueleret, Lydie
/ TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
/ FILE REFERENCE: 31. US3. CIP
/ CURRENT APPLICATION NUMBER: US/09/663, 600A
/ PRIOR APPLICATION NUMBER: 09/191, 997
/ PRIOR FILING DATE: 1998-11-13
/ PRIOR APPLICATION NUMBER: 60/066, 677
/ PRIOR FILING DATE: 1997-11-13
/ PRIOR APPLICATION NUMBER: 60/069, 957
/ PRIOR FILING DATE: 1997-12-17
/ PRIOR APPLICATION NUMBER: 60/074, 121
/ PRIOR FILING DATE: 1998-02-09
/ PRIOR APPLICATION NUMBER: 60/081, 563
/ PRIOR FILING DATE: 1998-04-13
/ PRIOR APPLICATION NUMBER: 60/096, 116
/ PRIOR FILING DATE: 1998-08-10
/ PRIOR APPLICATION NUMBER: 60/099, 273
/ PRIOR FILING DATE: 1998-09-04
/ NUMBER OF SEQ ID NOS: 229
/ SOFTWARE: Patent.pm
/ SEQ ID NO 149
/ LENGTH: 522
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: polyA_site
/ LOCATION: 512..522
/ NAME/KEY: misc_feature
/ LOCATION: 11
/ OTHER INFORMATION: n=a, g, c or t
US-09-663-600A-149

Alignment Scores:
Pred. No.: 1.98 Length: 522
Score: 62.00 Matches: 8
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 72.73% Mismatches: 2
Query Match: 73.81% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-663-600A-149 (1-522)
QY 1 HisAenHisAenHisAenHisAenHis 11
Db 441 CACTACATACACCAACCAATTCACATATCAT 473

RESULT 56
US-08-998-416-915
/ Sequence 915, Application US/08998416
/ Patent No. 6239264
/ GENERAL INFORMATION:
/ APPLICANT: Philippson, Peter
/ APPLICANT: Pohlmann, Rainer
/ APPLICANT: Steiner, Sabine
/ APPLICANT: Mohr, Christine
/ APPLICANT: Wendland, Jürgen
/ APPLICANT: Knechtel, Philipp
/ APPLICANT: Reibschung, Corinne
/ TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPHII
/ NUMBER OF SEQUENCES: 1152
/ AND USES THEREOF
/
CORRESPONDENCE ADDRESS:
/ ADDRESSER: No. 6239264artis Corporation
/ STREET: 3054 Cornwallis Road
/ CITY: Research Triangle Park
/ STATE: No. 6239264th Carolina
/ COUNTRY: USA
/ ZIP: 27709
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/998,416
/ FILING DATE: 24-DEC-1997
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: CH 0016/97
/ FILING DATE: 31-DEC-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Meiser, J. Timothy
/ REGISTRATION NUMBER: 38, 241
/ REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 919-541-8587
/ TELEFAX: 919-541-8689
/ INFORMATION FOR SEQ ID NO: 915:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 688 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ ORIGINAL SOURCE:
/ ORGANISM: PAG1570RP
US-08-998-416-915

Alignment Scores:
Pred. No.: 2.65 Length: 688
Score: 62.00 Matches: 10
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 2
Query Match: 73.81% Indels: 0
DB: 3 Gaps: 0

US-09-858-332G-15 (1-12) x US-08-998-416-915 (1-688)
QY 1 HisAenHisAenHisAenHisAenHisAen 12
Db 479 CACAACACCAACCAACCAATTCATGAACACCAT 514

RESULT 57
US-09-248-796A-7148
/ Sequence 7148, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstock et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR APPLICATION NUMBER: 1999-02-12
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 7148
/ LENGTH: 768
/ TYPE: DNA
/ ORGANISM: Candida albicans
/ FEATURE:

```

LOCATION: (749)
OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknd
US-09-248-796A-7148

Alignment Scores:
Pred. No.: 2.97 Length: 768
Score: 62.00 Matches: 9
Percent Similarity: 90.00% Conservative: 0
Best Local Similarity: 90.00% Mismatches: 1
Query Match: 73.81% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-7148 (1-768)

QY 2 AasnHisAsnHisAsnHisAsnHis 11
DB 263 AACCAACACCAACCAACATCCGACAC 292

RESULT 58
US-09-248-796A-713
Sequence 713, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 713
LENGTH: 1035
TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-713

Alignment Scores:
Pred. No.: 4.05 Length: 1035
Score: 62.00 Matches: 9
Percent Similarity: 83.33% Conservative: 1
Best Local Similarity: 75.00% Mismatches: 2
Query Match: 73.81% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-713 (1-1035)

QY 1 HisAsnHisAsnHisAsnHisAsnHis 12
DB 851 AACATCATCATCATCATCATCATCAGAGAGAC 886

RESULT 59
US-09-557-884-1/c
Sequence 1, Application US/09557884
Patent No. 6506581
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS V6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Alignment Scores:
Pred. No.: 1.34e+04 Length: 1830121
Score: 61.00 Matches: 8
Percent Similarity: 81.82% Conservative: 1
Best Local Similarity: 72.73% Mismatches: 2
Query Match: 72.62% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-557-884-1 (1-1830121)

QY 1 HisAsnHisAsnHisAsnHisAsnHis 11
DB 745469 CATCATCATGACATATCATCATCATATCATC 745437

RESULT 60
US-09-643-990A-1/c
Sequence 1, Application US/09643990A
Patent No. 6528289
GENERAL INFORMATION:
APPLICANT: Robert D. Fleischmann
Mark D. Adams
Owen White
Hamilton O. Smith
J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS V6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07

APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB186P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-8439
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Alignment Scores:
Pred. No.: 1,34e+04 Length: 1830121
Score: 61.00 Matches: 8
Percent Similarity: 81.82% Conservative: 1
Best Local Similarity: 72.73% Mismatches: 2
Query Match: 72.62% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-643-990A-1 (1-1830121)

OY 1 HisaenHisaenHisaenHisaenHis 11
DB 745469 CATCATCGAAGCATATCATCGACATATCATC 745437

RESULT 61
US-09-248-796A-4718
; Sequence 4718, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4718
; LENGTH: 858
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-4718

Alignment Scores:
Pred. No.: 6.06 Length: 858
Score: 60.00 Matches: 8
Percent Similarity: 83.33% Conservative: 2
Best Local Similarity: 66.67% Mismatches: 2
Query Match: 71.43% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-4718 (1-858)

OY 1 HisaenHisaenHisaenHisaenHis 12
DB 488 CACACACACCGCGCACCAACCAACCAACCAAC 523

RESULT 62
US-09-248-796A-3453
; Sequence 3453, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al

APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 3453
LENGTH: 888
TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-3453

Alignment Scores:
Pred. No.: 6.29 Length: 888
Score: 60.00 Matches: 8
Percent Similarity: 90.91% Conservative: 2
Best Local Similarity: 72.73% Mismatches: 1
Query Match: 71.43% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-3453 (1-888)

OY 1 HisaenHisaenHisaenHisaenHis 11
DB 161 CACACCGCGCACCAACCAACCAACCAATCATC 193

RESULT 63
US-09-248-796A-11317
; Sequence 11317, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 11317
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-11317

Alignment Scores:
Pred. No.: 7.53 Length: 1056
Score: 60.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 71.43% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-11317 (1-1056)

OY 1 HisaenHisaenHisaenHisaenHis 12
DB 795 AACCAACCAACCAACCAACCAACCAACCAAC 830

RESULT 64
US-09-248-796A-7374
; Sequence 7374, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
 FILE REFERENCE: 107196.132
 CURRENT APPLICATION NUMBER: US/09/248, 796A
 PRIOR FILING DATE: 1999-02-12
 PRIOR APPLICATION NUMBER: US 60/074, 725
 PRIOR FILING DATE: 1998-02-13
 PRIOR APPLICATION NUMBER: US 60/096, 409
 PRIOR FILING DATE: 1998-08-13
 NUMBER OF SEQ ID NOS: 28208
 SEQ ID NO 7374
 LENGTH: 1191
 TYPE: DNA
 ORGANISM: Candida albicans
 US-09-248-796A-7374

Alignment Scores:
 Pred. NO.: 8.54 Length: 1191
 Score: 60.00 Matches: 9
 Percent Similarity: 90.00% Conservative: 0
 Best Local Similarity: 90.00% Mismatches: 1
 Query Match: 71.43% Indels: 0
 DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-7374 (1-1191)

Oy 2 Aaahisaaahhisaaahhisaaahhis 11
 Db 893 AACCAATCCCAACCAATCACAATCAC 922

RESULT 65
 US-08-532-814-1/C
 Sequence 1, Application US/08532814
 Patent No. 5906817
 GENERAL INFORMATION:
 APPLICANT: MOULLIER, PHILIPPE
 APPLICANT: DANOS, OLIVIER
 APPLICANT: HEARD, JEAN-MICHEL
 APPLICANT: FERRY, NICOLAS
 TITLE OF INVENTION: BIOCOMPATIBLE IMPLANT FOR THE EXPRESSION
 TITLE OF INVENTION: AND IN VIVO SECRETION OF A THERAPEUTIC SUBSTANCE
 NUMBER OF SEQUENCES: 1
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
 CITY: ARLINGTON
 STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/532,814
 FILING DATE: 19-JAN-1996
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/FR94/00456
 FILING DATE: 21-APR-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 93/04700
 FILING DATE: 21-APR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 93/09185
 FILING DATE: 26-JUL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 660-105-0 PCT
 TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 8387 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-532-814-1

Alignment Scores:
Pred. No.:      65.5      Length:      8387
Score:          60.00     Matches:      8
Percent Similarity: 72.73% Conservative: 0
Best Local Similarity: 72.73% Mismatches:  3
Query Match:      71.43%  Indels:      0
DB:              2      Gaps:      0

US-09-858-3326-15 (1-12) x US-08-532-814-1 (1-8387)

Cy      1 HisasnhiasnHisasnhiasnHisasnhis 11
          |||||  |||  |||||  |||
Db      3437 CACAAATCACACACACAAACACATCACACACAC 3405

RESULT 66
US-09-225-509-1/c
; Sequence 1, Application US/09225509
; Patent No. 6326195
; GENERAL INFORMATION:
; APPLICANT: MOULIER, Phillipe
; APPLICANT: DANOS, Olivier
; APPLICANT: HEARD, Jean-Michel
; APPLICANT: FERRY, Nicholas
; TITLE OF INVENTION: BIOCOMPATIBLE IMPLANT FOR THE EXPRESSION AND IN VIVO
; TITLE OF INVENTION: SECRETION OF A THERAPEUTIC SUBSTANCE
; FILE REFERENCE: 0660-0145-0D1V
; CURRENT APPLICATION NUMBER: US/09/225,509
; CURRENT FILING DATE: 1999-01-06
; EARLIER APPLICATION NUMBER: PCT/FR94/00456
; EARLIER FILING DATE: 1994-04-21
; EARLIER APPLICATION NUMBER: 09/523,814
; EARLIER FILING DATE: 1996-01-19
; EARLIER APPLICATION NUMBER: FR 93/04700
; EARLIER FILING DATE: 1993-04-21
; EARLIER APPLICATION NUMBER: FR 93/09185
; EARLIER FILING DATE: 1993-07-26
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8388
; TYPE: DNA
; ORGANISM: mus musculus, Mo-Mulv, and other
US-09-225-509-1

Alignment Scores:
Pred. No.:      65.5      Length:      8388
Score:          60.00     Matches:      8
Percent Similarity: 72.73% Conservative: 0
Best Local Similarity: 72.73% Mismatches:  3
Query Match:      71.43%  Indels:      0
DB:              3      Gaps:      0

US-09-858-3326-15 (1-12) x US-09-225-509-1 (1-8388)

Cy      1 HisasnhiasnHisasnhiasnHisasnhis 11
          |||||  |||  |||||  |||
Db      3437 CACAAATCACACACACAAACACATCACACACAC 3405

RESULT 67
US-09-248-796A-14079/c
; Sequence 14079, Application US/09248796A
; Patent No. 6747137

```

```
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 14079
/ LENGTH: 192
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-14079

Alignment Scores:
Pred. No.: 1 72
Score: 59.00
Percent Similarity: 100.00%
Best Local Similarity: 58.33%
Query Match: 70.24%
DB: 4
Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-14079 (1-192)
QY
Db 1 H1SA8NH1SA8NH1SA8NH1SA8NH1SA8N 12
129 CATGACACACACACACACACACACACACAC 94

RESULT 68
US-09-248-796A-11235
/ Sequence 11235, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 11235
/ LENGTH: 228
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-11235

Alignment Scores:
Pred. No.: 2 05
Score: 59.00
Percent Similarity: 100.00%
Best Local Similarity: 58.33%
Query Match: 70.24%
DB: 4
Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-11235 (1-228)
QY
Db 1 H1SA8NH1SA8NH1SA8NH1SA8NH1SA8N 12
143 CACACACACACACACACACACACACACACAT 178

RESULT 69
US-09-248-796A-2784
/ Sequence 2784, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBIC
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 2784
/ LENGTH: 528
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-2784

Alignment Scores:
Pred. No.: 4 93
Score: 59.00
Percent Similarity: 100.00%
Best Local Similarity: 58.33%
Query Match: 70.24%
DB: 4
Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-2784 (1-528)
QY
Db 1 H1SA8NH1SA8NH1SA8NH1SA8NH1SA8N 12
310 CATGACACACACACACACACATCATCATCATC 345

RESULT 70
US-09-248-796A-9354
/ Sequence 9354, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 9354
/ LENGTH: 528
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-9354

Alignment Scores:
Pred. No.: 4 93
Score: 59.00
Percent Similarity: 73.33%
Best Local Similarity: 66.67%
Query Match: 70.24%
DB: 4
Gaps: 1

US-09-858-332G-15 (1-12) x US-09-248-796A-9354 (1-528)
QY
Db 2 A8NH1SA8N-----H1SA8NH1SA8NH1SA8N 12
217 AACATAACAACTCCACATCATCATCATATATAT 261

RESULT 71
US-09-949-016-43297
/ Sequence 43297, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
```

```

? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: CL001307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 43297
? LENGTH: 601
? TYPE: DNA
? ORGANISM: Human
? US-09-949-016-43297

```

Pred. No.:	5, 64	Length:	601
Score:	59.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	5
Best Local Similarity:	58.33%	Mismatches:	0
Query Match:	70.24%	Indels:	0
DB:	4	Gaps:	0

US-09-858-332G-15 (1-12) X US-09-949-016-43297 (1-601)

```
QY      1  HISSNTHISSNTHISSNTHISSNTHISSNTHISSN 12
          |||::|||::|||::|||::|||::|||::|||
Db      268 CATCATCATCATCATCATCATCATCATCATCAAC 303
```

RESULT 72

Sequence 43526, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43526
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-43526

Pred. No.:	5.64	Length:	601
Score:	59.00	Matches:	7
Percent Similarity:	100.00*	Conservative:	5
Best local Similarity:	58.33%	Mismatches:	0
Query Match:	70.24%	Indels:	0
DB:	4	Gaps:	0

US-09-858-332G-15 (1-12) x US-09-949-016-43526 (1-601)

```

QY      1  HIASNHIASNHIASNHIASNHIASNHIASN 12
        ||::||::||::||::||::||::||
Db      268 CATCATCATCATCATCATCATCATCATTAAC 303

```

RESULT 73

```

US-09-949-016-43755
: Sequence 43755, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001307
: CURRENT APPLICATION NUMBER: US/09/949, 016
: PRIOR FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 43755
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-43755

```

Pred. No.:	5.64	Length:	601
Score:	59.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	5
Best Local Similarity:	58.33%	Mismatches:	0
Query Match:	70.24%	Indels:	0
DB:	4	Gaps:	0

US-09-09-858-332G-15 (1-12) x US-09-949-016-43755 (1-601)

```

Qy      1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
          |||::|||::|||::|||::|||::|||
Db      268 CATCATCATCATCATCATCATCATCATCATAC 303

```

RESULT 74

```

Sequence 49768, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949.016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 49768

```

```

!      LENGTH: 302
!      TYPE: DNA
!      ORGANISM: Human
US-09-949-016-49768
Alignment Scores:

```

Percent Similarity:

Query Match:	70.24%	Indels:	0
DB:	4	Gaps:	0

US-09-858-332G-15 (1-12) x US-09-949-016-49768 (1-601)

RESULT 83

US-09-248-796A-6566
; Sequence 6566, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 6566
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-6566

Alignment Scores:

Pred. No.: 11.7 Length: 1212
Score: 59.00 Matches: 7
Percent Similarity: 100.00% Conservative: 5
Best Local Similarity: 58.33% Mismatches: 0
Query Match: 70.24% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-248-796A-6566 (1-1212)

QY 1 Hisaenhisaenhisaenhisaenhisaen 12

Db 325 CACCATCATCACCACCATATCATCATCATCA 360

RESULT 84

US-09-900-708-1
; Sequence 1, Application US/09900708
; Patent No. 6762342
; GENERAL INFORMATION:
; APPLICANT: Allen, Keith D.
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING INTESTINAL
; FILE REFERENCE: R-733
; CURRENT APPLICATION NUMBER: US/09/900,708
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US 60/216,476
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: US 60/221,489
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PaSeq For Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5293
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-900-708-1

Alignment Scores:

Pred. No.: 54.6 Length: 5293
Score: 59.00 Matches: 7
Percent Similarity: 100.00% Conservative: 5
Best Local Similarity: 58.33% Mismatches: 0
Query Match: 70.24% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-900-708-1 (1-5293)

QY 1 Hisaenhisaenhisaenhisaenhisaen 12

Db 3268 CATCACACACACACACACACACACACACACACACACACAC 3303

RESULT 85

US-09-627-465B-1/c
; Sequence 1, Application US/09627465B
; Patent No. 6737519
; GENERAL INFORMATION:
; APPLICANT: KEITH, TIM
; APPLICANT: LITTLE, RANDALL D.
; APPLICANT: BERDEWEGH, PAUL VAN
; APPLICANT: DUPUIS, JOSEF
; APPLICANT: DEL MASTRO, RICHARD L.
; APPLICANT: SIMON, JASON
; APPLICANT: ALLEN, KRISTINA
; APPLICANT: PANDIT, SUNIL
; TITLE OF INVENTION: NOVEL HUMAN GENES RELATING TO RESPIRATORY DISEASES AND OBESITY
; FILE REFERENCE: 2976-4037
; CURRENT APPLICATION NUMBER: US/09/627,465B
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/211,749
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 60/146,336
; PRIOR FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver 2.1
; SEQ ID NO 1
; LENGTH: 10304
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (267)
; OTHER INFORMATION: a, t, c or g
US-09-627-465B-1

Alignment Scores:

Pred. No.: 110 Length: 10304
Score: 59.00 Matches: 7
Percent Similarity: 100.00% Conservative: 5
Best Local Similarity: 58.33% Mismatches: 0
Query Match: 70.24% Indels: 0
DB: 4 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-627-465B-1 (1-10304)

QY 1 Hisaenhisaenhisaenhisaenhisaen 12

Db 9955 CATCACACCATCATCATCATCATCATCATCATCAT 9920

RESULT 86

US-09-356-952-12/c
; Sequence 12, Application US/09356952
; Patent No. 6117663
; GENERAL INFORMATION:
; APPLICANT: Borjuck-Sjodin, Ann
; APPLICANT: Margarit, S. M.
; APPLICANT: Bor-Sogit, Dafna
; APPLICANT: Cole, Philip
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
; FILE REFERENCE: 600-1-228N
; CURRENT APPLICATION NUMBER: US/09/356,952
; PRIOR FILING DATE: 1999-07-19
; EARLIER APPLICATION NUMBER: 60/093,631
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 43676
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-356-952-12

QY 1 Hisaenhisaenhisaenhisaenhisaen 12

Db 9955 CATCACACCATCATCATCATCATCATCATCATCAT 9920

Db 189044 CACGATCATACCATATGATATCATGTAACAT 189076

RESULT 94
US-08-097-829-7/c
Sequence 7, Application US/08097829
Patent No. 5498831
GENERAL INFORMATION:
APPLICANT: Burgess, Diane G.
APPLICANT: Dooner, Hugo K.
TITLE OF INVENTION: Pea ADP-Glucose Pyrophosphorylase Subunit
TITLE OF INVENTION: Genes and Their Uses
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Hourie and Crew
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,829
FILING DATE: 19930723
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 12176-35
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-097-829-7

Alignment Scores:
Pred. No.: 3.58 Length: 292
Score: 58.00 Matches: 8
Percent Similarity: 81.82% Conservative: 1
Best Local Similarity: 72.73% Mismatches: 2
Query Match: 69.05% Indels: 0
Gaps: 0

US-09-858-332g-15 (1-12) x US-08-097-829-7 (1-292)

QY 1 HisAsnHisAsnHisAsnHisAsnHis 11
Db 218 TACACTCAAAATCATATATCATATATCAT 186

RESULT 95
US-08-577-403-7/c
Sequence 7, Application US/08577403
Patent No. 5771693
GENERAL INFORMATION:
APPLICANT: Burgess, Diane G.
APPLICANT: Dooner, Hugo K.
TITLE OF INVENTION: Pea ADP-Glucose Pyrophosphorylase Subunit
TITLE OF INVENTION: Genes and Their Uses
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Hourie and Crew
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California

COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/577,403
FILING DATE: 22-DEC-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/097,829
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 12176-35
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-577-403-7

Alignment Scores:
Pred. No.: 3.58 Length: 292
Score: 58.00 Matches: 8
Percent Similarity: 81.82% Conservative: 1
Best Local Similarity: 72.73% Mismatches: 2
Query Match: 69.05% Indels: 0
Gaps: 0

US-09-858-332g-15 (1-12) x US-08-577-403-7 (1-292)

QY 1 HisAsnHisAsnHisAsnHisAsnHis 11
Db 218 TACACTCAAAATCATATATCATATATCAT 186

RESULT 96
US-09-248-796A-9887
Sequence 9887, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCES: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 9887
LENGTH: 435
TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-9887

Alignment Scores:
Pred. No.: 5.43 Length: 435
Score: 58.00 Matches: 7
Percent Similarity: 91.67% Conservative: 4
Best Local Similarity: 58.33% Mismatches: 1
Query Match: 69.05% Indels: 0
Gaps: 0

Alignment Scores:

Pred. No.:	21.8	Length:	1644
Score:	58.00	Matches:	8
Percent Similarity:	90.91%	Conservative:	2
Best Local Similarity:	72.73%	Mismatches:	1
Query Match:	69.05%	Indels:	0
DB:	4	Gaps:	0

US-09-858-332g-15 (1-12) x US-09-248-796A-858 (1-1644)

QY	1	HisAsnHisAsnHisAsnHisAsnHisAsnHis	11
		::: :::	
Db	92	CACTACACACACACACACACACACACACACAC	124

Search completed: October 4, 2005, 13:27:56
Job time : 474 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: October 4, 2005, 11:50:22 ; Search time 626 Seconds
(without alignments)
132.358 Million cell updates/sec

Title: US-09-858-332G-15
Perfect score: 84
Sequence: 1 HNNHNNHNNHNN 12

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 7442561 segs, 3452328358 residues
Total number of hits satisfying chosen parameters: 14885122

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+.p2n.model -DEV=xlp
-Q=/cgn2_1/USFTO.spool_p/US09858332/runat_04102005_122258_4354/app_query.fasta.1.199
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=p2n.rmpb -MINMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS-human40.csi -LIST=100 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=100 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=2000000000
-USER=US09858332.@CGN_1.1.723.@runat_04102005_122258_4354 -NCPUS=6 -ICU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	84	100.0	36	9	US-09-858-332-17	Sequence 17, Appl
2	84	100.0	42	14	US-10-055-794-13	Sequence 13, Appl
3	84	100.0	100	14	US-10-055-794-11	Sequence 11, Appl
4	84	100.0	115	16	US-10-125-594-3	Sequence 3, Appl
5	84	100.0	115	17	US-10-429-979-3	Sequence 3, Appl
6	84	100.0	519	17	US-10-191-803-242	Sequence 242, Appl
7	84	100.0	540	21	US-10-486-706-55	Sequence 55, Appl
8	84	100.0	1552	19	US-10-437-963-5758	Sequence 5758, Ap
9	84	100.0	1698	19	US-10-451-467A-193	Sequence 193, App
10	84	100.0	2258	17	US-10-429-979-6	Sequence 6, Appl
11	84	100.0	3916	24	US-11-003-103A-5	Sequence 5, Appl
12	84	100.0	4938	14	US-10-055-794-3	Sequence 1, Appl
13	84	100.0	6525	14	US-10-055-794-4	Sequence 3, Appl
14	84	100.0	7487	14	US-10-055-794-4	Sequence 4, Appl
15	78	92.9	139	16	US-10-125-594-5	Sequence 5, Appl
16	78	92.9	139	17	US-10-429-979-5	Sequence 5, Appl
17	78	92.9	205	16	US-10-125-594-2	Sequence 2, Appl
18	78	92.9	205	17	US-10-429-979-2	Sequence 2, Appl
19	78	92.9	587	16	US-10-125-594-1	Sequence 1, Appl
20	78	92.9	587	17	US-10-429-979-1	Sequence 1, Appl
21	78	92.9	2580	19	US-10-793-639-261	Sequence 261, App
22	78	92.9	3455	9	US-09-826-752-3	Sequence 3, Appl
23	78	92.9	3455	21	US-10-912-434-3	Sequence 3, Appl
24	77	91.7	537	19	US-10-021-323-11036	Sequence 11036, A
25	77	91.7	698	18	US-10-424-599-51499	Sequence 51499, A
26	77	91.7	16489	9	US-09-764-868-1483	Sequence 1483, Ap
27	77	91.7	54718	20	US-10-719-993-7038	Sequence 7038, Ap
28	76	90.5	537	14	US-10-154-971-21	Sequence 21, Appl
29	76	90.5	912	14	US-10-154-971-23	Sequence 23, Appl
30	76	90.5	1154	10	US-09-737-476B-54	Sequence 54, Appl
31	76	90.5	1172	10	US-09-737-476B-56	Sequence 56, Appl
32	76	90.5	1521	16	US-10-033-585-6045	Sequence 6045, Ap
33	76	90.5	1521	21	US-10-882-104-7	Sequence 7, Appl
34	76	90.5	4565	9	US-09-794-384A-3	Sequence 3, Appl
35	76	90.5	4565	13	US-10-138-221-8	Sequence 8, Appl
36	76	90.5	4565	20	US-10-801-550-3	Sequence 3, Appl
37	76	90.5	4798	19	US-10-425-115-19825	Sequence 19825, A
38	76	90.5	4810	13	US-10-138-221-6	Sequence 6, Appl
39	74	88.1	1921	19	US-10-437-963-45568	Sequence 45568, A
40	74	85.7	3942	19	US-10-451-467A-545	Sequence 545, App
41	72	85.7	5826	15	US-10-311-455-536	Sequence 536, App
42	72	85.7	6363	15	US-10-311-455-1603	Sequence 1603, Ap
43	72	85.7	3673778	16	US-10-312-841-2	Sequence 2, Appl
44	71	84.5	3186778	17	US-10-027-632-174961	Sequence 174961, Sequence 1, Appl
45	71	84.5	3186778	13	US-10-027-632-174961	Sequence 1, Appl
46	70	83.3	2556	9	US-09-817-310-1	Sequence 641, App
47	70	83.3	24076	19	US-10-451-467A-641	Sequence 179264, Sequence 19264, Sequence 107879, Sequence 57663, A
48	70	83.3	1223197	17	US-10-027-632-179264	Sequence 50046, A
49	69	82.1	550	18	US-10-424-599-107879	Sequence 96040, A
50	69	82.1	676	18	US-10-437-963-50040	Sequence 19761, A
51	69	82.1	1385	19	US-10-425-115-19761	Sequence 97126, A
52	69	82.1	3855	24	US-11-097-143-42397	Sequence 996, App
53	69	82.1	10907	24	US-11-097-143-42397	Sequence 1275, App
54	69	82.1	10907	24	US-11-097-143-42397	Sequence 535, App
55	68.5	81.5	2094	24	US-11-097-143-21194	Sequence 1604, Ap
56	68.5	81.5	4569	24	US-11-097-143-21193	Sequence 24139, A
57	67	79.8	4779	24	US-11-097-143-24139	Sequence 96040, A
58	66	78.6	331	18	US-10-424-599-96040	Sequence 97126, A
59	66	78.6	449	20	US-10-425-115-19761	Sequence 996, App
60	66	78.6	1335	18	US-10-424-599-97126	Sequence 1275, App
61	66	78.6	1335	17	US-10-374-780A-996	Sequence 535, App
62	66	78.6	1338	18	US-10-412-699B-1275	Sequence 1604, Ap
63	66	78.6	5826	15	US-10-311-455-535	Sequence 1604, Ap
64	66	78.6	6236	15	US-10-311-455-1604	Sequence 1604, Ap

65 66 78.6 28360 24 US-11-097-143-6454
66 66 78.6 3673778 16 US-10-312-841-1
67 65 77.4 334 9 US-09-728-445-386
68 65 77.4 334 22 US-10-964-549-386
69 65 77.4 700 10 US-09-327-750-55
70 65 77.4 927 24 US-11-097-143-38234
71 65 77.4 1958 24 US-11-097-143-38233
72 65 77.4 13033 24 US-11-097-143-18322
73 65 77.4 684707 17 US-10-398-221-9
74 65 77.4 3011208 17 US-10-398-221-2058
75 64 76.2 65 16 US-10-032-585-1299
76 64 76.2 148 9 US-09-864-761-19229
77 64 76.2 381 9 US-09-864-761-2500
78 64 76.2 399 18 US-10-424-599-51775
79 64 76.2 413 19 US-10-674-124A-3759
80 64 76.2 444 9 US-09-770-444-818
81 64 76.2 479 9 US-09-864-761-768
82 64 76.2 515 10 US-09-770-961-127
83 64 76.2 559 16 US-10-029-386-3710
84 64 76.2 609 19 US-10-767-701-26149
85 64 76.2 629 18 US-10-424-599-109460
86 64 76.2 830 9 US-09-864-761-19531
87 64 76.2 1356 17 US-10-282-122A-11026
88 64 76.2 1457 18 US-10-424-599-37489
89 64 76.2 1898 24 US-11-097-143-22082
90 64 76.2 2257 20 US-10-739-930-229
91 64 76.2 12396 24 US-11-097-143-22081
92 64 76.2 21379 24 US-11-097-143-42484
93 64 76.2 21409 24 US-11-097-143-42451
94 64 76.2 21449 24 US-11-097-143-2977
95 64 76.2 21451 24 US-11-097-143-28753
96 64 76.2 21451 24 US-11-097-143-28756
97 64 76.2 126192 13 US-10-672-787-21
98 64 76.2 126192 13 US-10-087-192-1459
99 64 76.2 135259 18 US-10-260-425-1585
C 100 64 76.2 249487 13 US-10-026-188-3

ALIGNMENTS

Sequence 6454, App
Sequence 1, Appli
Sequence 386, App
Sequence 386, App
Sequence 55, Appl
Sequence 38233, A
Sequence 38233, A
Sequence 18322, A
Sequence 9, Appl
Sequence 2058, Ap
Sequence 1929, Ap
Sequence 19229, A
Sequence 2500, A
Sequence 51775, A
Sequence 3759, Ap
Sequence 818, App
Sequence 768, App
Sequence 127, App
Sequence 26149, A
Sequence 109460, A
Sequence 19531, A
Sequence 11026, A
Sequence 37489, A
Sequence 22082, A
Sequence 229, App
Sequence 22081, A
Sequence 42484, A
Sequence 42451, A
Sequence 2977, App
Sequence 28753, A
Sequence 28756, A
Sequence 21, Appl
Sequence 1459, App
Sequence 1585, App
Sequence 3, Appl

RESULT 1
US-09-858-332-17
Sequence 17, Application US/09858332
Patent No. US20020164718A1
GENERAL INFORMATION:
APPLICANT: Tchnaga, Grigory S.
TITLE OF INVENTION: Metal Ion Affinity Tags and Methods for
FILE REFERENCE: CLON056CIP
CURRENT APPLICATION NUMBER: US/09/858,332
PRIOR FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: 09/404,017
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: 60/101,867
PRIOR FILING DATE: 1998-09-25
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 36
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
US-09-858-332-17

Alignment Scores:

Pred. No.: 0.00072 Length: 36
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-858-332G-15 (1-12) x US-09-858-332-17 (1-36)

OY 1 H1S8NH1S8NH1S8NH1S8NH1S8NH1S8N 12
DB 1 CATATCATATCATATCATATCATATCATATCAAC 36

RESULT 2

US-10-055-794-13
Sequence 13, Application US/10055794
Publication No. US20030059900A1
GENERAL INFORMATION:
APPLICANT: Farmer, Alan Andrew
TITLE OF INVENTION: SEQUENCE SPECIFIC RECOMBINASE-BASED
METHODS FOR PRODUCING INTRON CONTAINING VECTORS AND
FILE REFERENCE: CLON-069
CURRENT APPLICATION NUMBER: US/10/055,794
PRIOR FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: 60/263,358
PRIOR FILING DATE: 2001-01-18
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 42
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: splice sequence
US-10-055-794-13

Alignment Scores:

Pred. No.: 0.000849 Length: 42
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-858-332G-15 (1-12) x US-10-055-794-13 (1-42)

OY 1 H1S8NH1S8NH1S8NH1S8NH1S8NH1S8N 12
DB 4 CATATCATATCATATCATATCATATCATATCAAC 39

RESULT 3

US-10-055-794-11
Sequence 11, Application US/10055794
Publication No. US20030059900A1
GENERAL INFORMATION:
APPLICANT: Farmer, Alan Andrew
TITLE OF INVENTION: SEQUENCE SPECIFIC RECOMBINASE-BASED
METHODS FOR PRODUCING INTRON CONTAINING VECTORS AND
FILE REFERENCE: CLON-069
CURRENT APPLICATION NUMBER: US/10/055,794
PRIOR FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: 60/263,358
PRIOR FILING DATE: 2001-01-18
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 100
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: splice sequence
US-10-055-794-11

Alignment Scores:

Pred. No.: 0.00215 Length: 100
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0

;; PRIOR APPLICATION NUMBER: US 60/311,343
;; PRIOR FILING DATE: 2001-08-13
;; NUMBER OF SEQ ID NOS: 461
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 55
;; LENGTH: 540
;; TYPE: DNA
;; ORGANISM: Rattus norvegicus
US-10-486-706-55

Alignment Scores:

Pred. No.:	0.0131	Length:	540
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0

US-09-858-332g-15 (1-12) x US-10-486-706-55 (1-540)

QY 1 Hisaenhisaenhisaenhisaenhisaen 12
Db 537 CATACCAACCAACCAACCAACCAACCAAT 502

RESULT 8

US-10-437-963-5758
; Sequence 5758, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437, 963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 5758
; LENGTH: 1652
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1652)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_12513C.1
US-10-437-963-5758

Alignment Scores:

Pred. No.:	0.0436	Length:	1652
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	19	Gaps:	0

US-09-858-332g-15 (1-12) x US-10-437-963-5758 (1-1652)

QY 1 Hisaenhisaenhisaenhisaenhisaen 12
Db 217 CACACCAACCAACCAACCAACCAAT 252

RESULT 9

US-10-451-467A-193
; Sequence 193, Application US/10451467A
; Publication No. US20040161840A1

;; GENERAL INFORMATION:
;; APPLICANT: CONTRERAS, ROLAND HENRI
;; APPLICANT: EBERHARDT, INES
;; APPLICANT: LUTTEN, WALTER HERMAN MARIA LOUIS
;; APPLICANT: REEKMAN, RIEKA JOSEPHINA
;; TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN
;; FILE REFERENCE: JAB-1667
;; CURRENT APPLICATION NUMBER: US/10/451,467A
;; CURRENT FILING DATE: 2003-06-19
;; PRIOR APPLICATION NUMBER: EP 00870318.3
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: EP 01870002.1
;; PRIOR FILING DATE: 2001-01-04
;; PRIOR APPLICATION NUMBER: EP 01870003.9
;; PRIOR FILING DATE: 2001-01-09
;; NUMBER OF SEQ ID NOS: 732
;; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 193
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-451-467A-193

Alignment Scores:

Pred. No.:	0.0506	Length:	1898
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	19	Gaps:	0

US-09-858-332g-15 (1-12) x US-10-451-467A-193 (1-1898)

QY 1 Hisaenhisaenhisaenhisaenhisaen 12
Db 195 CATATCATATCATATCATATCATATCATAT 230

RESULT 10

US-10-429-979-6/C
; Sequence 6, Application US/10429979
; Publication No. US20030224433A1
; GENERAL INFORMATION:
; APPLICANT: Kaushik, Azad
; APPLICANT: Saini, Surinder Singh
; APPLICANT: Shojael, Farbod
; TITLE OF INVENTION: Bovine D-genes
; FILE REFERENCE: 12837-5
; CURRENT APPLICATION NUMBER: US/10/429,979
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: US 10/125,594
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/284,899
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 2258
; TYPE: DNA
; ORGANISM: Bovine
US-10-429-979-6

Alignment Scores:

Pred. No.:	0.061	Length:	2258
Score:	84.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	17	Gaps:	0

US-09-858-332g-15 (1-12) x US-10-429-979-6 (1-2258)

QY 1 Hisaenhisaenhisaenhisaenhisaen 12

Db 1532 CATATACATATACATATACCAACCAATACATACAC 1497

RESULT 11

US-11-003-103A-5
; Sequence 5, Application US/11003103A
; Publication No. US20050152888A1
; GENERAL INFORMATION:
; APPLICANT: Church, George
; APPLICANT: Janse, Daniel
; TITLE OF INVENTION: Targeted Protein Degradation
; FILE REFERENCE: 10498-00076
; CURRENT APPLICATION NUMBER: US/11/003,103A
; CURRENT FILING DATE: 2004-12-03
; PRIOR APPLICATION NUMBER: 60/526,490
; PRIOR FILING DATE: 2003-12-03
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5
; LENGTH: 3916
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid PROTEC133-TOR-link-HIS3HA
US-11-003-103A-5

Alignment Scores:

Pred. No.: 0.11 Length: 3916
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-09-858-332g-15 (1-12) x US-11-003-103A-5 (1-3916)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 114 CATATATCATATATCATATATCATATATCATATATCATATAC 149

RESULT 12

US-10-055-794-1
; Sequence 1, Application US/10055794
; Publication No. US20030059900A1
; GENERAL INFORMATION:
; APPLICANT: Farmer, Alan Andrew
; TITLE OF INVENTION: SEQUENCE SPECIFIC RECOMBINASE-BASED
; TITLE OF INVENTION: METHODS FOR PRODUCING INTRON CONTAINING VECTORS AND
; TITLE OF INVENTION: COMPOSITIONS FOR USE IN PRACTICING THE SAME
; FILE REFERENCE: CLON-069
; CURRENT APPLICATION NUMBER: US/10/055,794
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/263,358
; PRIOR FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4938
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Vector
US-10-055-794-1

Alignment Scores:

Pred. No.: 0.141 Length: 4938
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-858-332g-15 (1-12) x US-10-055-794-1 (1-4938)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 141 CATATATCATATATCATATATCATATATCATATATCATATAC 176

RESULT 13

US-10-055-794-3
; Sequence 3, Application US/10055794
; Publication No. US20030059900A1
; GENERAL INFORMATION:
; APPLICANT: Farmer, Alan Andrew
; TITLE OF INVENTION: SEQUENCE SPECIFIC RECOMBINASE-BASED
; TITLE OF INVENTION: METHODS FOR PRODUCING INTRON CONTAINING VECTORS AND
; TITLE OF INVENTION: COMPOSITIONS FOR USE IN PRACTICING THE SAME
; FILE REFERENCE: CLON-069
; CURRENT APPLICATION NUMBER: US/10/055,794
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/263,358
; PRIOR FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 6525
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Vector
US-10-055-794-3

Alignment Scores:

Pred. No.: 0.19 Length: 6525
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-858-332g-15 (1-12) x US-10-055-794-3 (1-6525)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

Db 1728 CATATATCATATATCATATATCATATATCATATATCATATAC 1763

RESULT 14

US-10-055-794-4
; Sequence 4, Application US/10055794
; Publication No. US20030059900A1
; GENERAL INFORMATION:
; APPLICANT: Farmer, Alan Andrew
; TITLE OF INVENTION: SEQUENCE SPECIFIC RECOMBINASE-BASED
; TITLE OF INVENTION: METHODS FOR PRODUCING INTRON CONTAINING VECTORS AND
; TITLE OF INVENTION: COMPOSITIONS FOR USE IN PRACTICING THE SAME
; FILE REFERENCE: CLON-069
; CURRENT APPLICATION NUMBER: US/10/055,794
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/263,358
; PRIOR FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 7487
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Vector
US-10-055-794-4

Alignment Scores:

Pred. No.: 0.221 Length: 7487
Score: 84.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

DB: 17 Gaps: 0
US-09-858-332g-15 (1-12) x US-10-429-979-2 (1-205)
QY 1 H1sAsnH1sAsnH1sAsnH1sAsnH1sAsn 12
:::|||||
DB 86 TACAAACCATTAACCATTAACCAACCATTAACCATTAAC 51
RESULT 19
US-10-125-594-1/C
; Sequence 1, Application US/10125594
; Publication No. US20030170646A1
; GENERAL INFORMATION:
; APPLICANT: Kaushik, Azad
; APPLICANT: Saini, Surinder Singh
; TITLE OF INVENTION: No. US20030170646A1 Bovine VDJ Casette, BPH1, Suitable for Ar
; FILE REFERENCE: 12837-4
; CURRENT APPLICATION NUMBER: US/10/125,594
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/284,899
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 587
; TYPE: DNA
; ORGANISM: Bovine
US-10-125-594-1
Alignment Scores:
Pred. No.: 0.0897 Length: 587
Score: 78.00 Matches: 11
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 91.67% Mismatches: 0
Query Match: 92.86% Indels: 0
Gaps: 0
DB: 16
US-09-858-332g-15 (1-12) x US-10-125-594-1 (1-587)
QY 1 H1sAsnH1sAsnH1sAsnH1sAsnH1sAsn 12
:::|||||
DB 417 TACAAACCATTAACCATTAACCAACCATTAACCATTAAC 382
RESULT 20
US-10-429-979-1/C
; Sequence 1, Application US/10429979
; Publication No. US2003022443A1
; GENERAL INFORMATION:
; APPLICANT: Kaushik, Azad
; APPLICANT: Saini, Surinder Singh
; APPLICANT: Shojiei, Farbod
; TITLE OF INVENTION: Bovine D-Genes
; FILE REFERENCE: 12837-5
; CURRENT APPLICATION NUMBER: US/10/429,979
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: US 10/125,594
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/284,899
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 587
; TYPE: DNA
; ORGANISM: Bovine
US-10-429-979-1
Alignment Scores:
Pred. No.: 0.0897 Length: 587
Score: 78.00 Matches: 11
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 91.67% Mismatches: 0
Query Match: 92.86% Indels: 0
Gaps: 0
DB: 0

DB: 17 Gaps: 0
US-09-858-332g-15 (1-12) x US-10-429-979-1 (1-587)
QY 1 H1sAsnH1sAsnH1sAsnH1sAsnH1sAsn 12
:::|||||
DB 417 TACAAACCATTAACCATTAACCAACCATTAACCATTAAC 382
RESULT 21
US-10-793-639-261
; Sequence 261, Application US/10793639
; Publication No. US2004019940A1
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasubramanian
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH STEROL SYNTHESIS AND METABOLISM
; FILE REFERENCE: 16516.075
; CURRENT APPLICATION NUMBER: US/10/793,639
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US/09/614,221A
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: US 60/142,981
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 261
; LENGTH: 2580
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-793-639-261
Alignment Scores:
Pred. No.: 0.439 Length: 2580
Score: 78.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.86% Indels: 0
Gaps: 0
DB: 19
US-09-858-332g-15 (1-12) x US-10-793-639-261 (1-2580)
QY 1 H1sAsnH1sAsnH1sAsnH1sAsnH1sAsn 11
:::|||||
DB 1900 CATTAATCATTAATCATTAATCATTAATCATTAATCATTAAC 1932
RESULT 22
US-09-826-752-3
; Sequence 3, Application US/09826752
; Patent No. US20010026930A1
; GENERAL INFORMATION:
; APPLICANT: Guarente, Leonard P.
; APPLICANT: Austriaco Jr., Nicanor
; APPLICANT: Claus, James J.
; APPLICANT: Cole, Francesca
; APPLICANT: Kennedy, Brian
; TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESECE IN
; TITLE OF INVENTION: YEAST
; FILE REFERENCE: 0050.1491-005
; CURRENT APPLICATION NUMBER: US/09/826,752
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 08/396,001
; PRIOR FILING DATE: 1995-02-28
; PRIOR APPLICATION NUMBER: PCT/US94/09351
; PRIOR FILING DATE: 1994-08-15
; PRIOR APPLICATION NUMBER: US 08/107,408
; PRIOR FILING DATE: 1993-08-16
; PRIOR APPLICATION NUMBER: US 09/323,433
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3455

US-09-858-332g-15 (1-12) x US-10-424-599-51499 (1-698)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

DB 688 CACAAATCAATCAATCAATCAATCAATCAATCAAT 653

RESULT 26

US-09-764-868-1483
; Sequence 1483, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1483
; LENGTH: 16489
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-1483

Alignment Scores:

Pred. No.:	4,36	Length:	16489
Score:	77.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	91.67%	Mismatches:	0
Query Match:	91.67%	Indels:	0
DB:	9	Gaps:	0

US-09-858-332g-15 (1-12) x US-09-764-868-1483 (1-16489)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

DB 7017 AATATCATATATCATATATCATATATCATATATCATAT 7052

RESULT 27

US-10-719-993-7038
; Sequence 7038, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7038
; LENGTH: 54718
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-7038

Alignment Scores:

Pred. No.:	15,8	Length:	54718
Score:	77.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	91.67%	Mismatches:	0
Query Match:	91.67%	Indels:	0
DB:	20	Gaps:	0

US-09-858-332g-15 (1-12) x US-10-719-993-7038 (1-54718)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

DB 48069 AATATCATATATCATATATCATATATCATATATCATAT 48104

RESULT 28

US-10-154-971-21

; Sequence 21, Application US/10154971

; Publication No. US20030088074A1

GENERAL INFORMATION:

APPLICANT: Hamers, Raymond

Muyldermans, Serge

TITLE OF INVENTION: VARIABLE FRAGMENTS OF IMMUNOGLOBULINS -

USE FOR THERAPEUTIC OR VETERINARY PURPOSES

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: SPENCER & FRANK

STREET: 1100 New York Avenue, N.W., Suite 300 East

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/154,971

FILING DATE: 28-May-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/945,244

FILING DATE: <Unknown>

APPLICATION NUMBER: PCT/EP96/01725

FILING DATE: 25-APR-1996

APPLICATION NUMBER: EP 95400932.0

FILING DATE: 25-APR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Gollin, Michael A.

REGISTRATION NUMBER: 31,957

REFERENCE/DOCKET NUMBER: GUPLA 0003

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-414-4000

TELEFAX: 202-414-4040

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 537 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 1..534

SEQUENCE DESCRIPTION: SEQ ID NO: 21:

US-10-154-971-21

Alignment Scores:

Pred. No.:	0,15	Length:	537
Score:	76.00	Matches:	11
Percent Similarity:	91.67%	Conservative:	0
Best Local Similarity:	91.67%	Mismatches:	1
Query Match:	90.48%	Indels:	0
DB:	14	Gaps:	0

US-09-858-332g-15 (1-12) x US-10-154-971-21 (1-537)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

DB 413 CACAAACCAACCAACCAACCAACCAACCAACCAAC 448

RESULT 29

US-10-154-971-23

; Sequence 23, Application US/10154971

; Publication No. US20030088074A1

GENERAL INFORMATION:

APPLICANT: Hamers, Raymond

Muyldermans, Serge
TITLE OF INVENTION: VARIABLE FRAGMENTS OF IMMUNOGLOBULINS -
USE FOR THERAPEUTIC OR VETERINARY PURPOSES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SPENCER & FRANK
STREET: 1100 New York Avenue, N.W., Suite 300 East
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/154,971
FILING DATE: 28-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/945,244
FILING DATE: <Unknown>
APPLICATION NUMBER: PCT/EP96/01725
FILING DATE: 25-APR-1996
APPLICATION NUMBER: EP 95400932.0
FILING DATE: 25-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Gollin, Michael A.
REGISTRATION NUMBER: 31,957
REFERENCE/DOCKET NUMBER: GUPLA 0003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-414-4000
TELEFAX: 202-414-4040
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 912 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..909
SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-10-154-971-23
Alignment Scores:
Pred. No.: 0.265
Score: 76.00
Percent Similarity: 91.67%
Best Local Similarity: 91.67%
Query Match: 90.48%
DB: 14
Gaps: 0
US-09-858-332g-15 (1-12) x US-10-154-971-23 (1-912)
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 413 CACACCAACCAACCAACCAACCAACCAACCAAC 448
RESULT 30
US-09-737-476B-54
Sequence 54, Application US/09737476B
Publication No. US20030208797A1
GENERAL INFORMATION:
APPLICANT: Leo G.J. FRENKEN
APPLICANT: Cornelis P.E. van der LOGT
APPLICANT: Vin-Min TEH
APPLICANT: Martine E. VERHOEYEN
APPLICANT: Joy E. WILKINSON
APPLICANT: Stephen A. JOBLING
TITLE OF INVENTION: Production of Antibodies

FILE REFERENCE: PAK/060113/0275850 - T7060C
CURRENT APPLICATION NUMBER: US/09/737,476B
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: EP 99310188.0
PRIOR FILING DATE: 1999-12-17
NUMBER OF SEQ ID NOS: 67
SOFTWARE: MS Word
SEQ ID NO 54
LENGTH: 1154
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HCV33-hinge-CH2-CH3
NAME/KEY: CDS
LOCATION: (3) .. (1136)
US-09-737-476B-54
Alignment Scores:
Pred. No.: 0.341
Score: 76.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 90.48%
DB: 10
Gaps: 0
US-09-858-332g-15 (1-12) x US-09-737-476B-54 (1-1154)
QY 2 AsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 406 AACCAACCAACCAACCAACCAACCAACCAAC 438
RESULT 31
US-09-737-476B-56
Sequence 56, Application US/09737476B
Publication No. US20030208797A1
GENERAL INFORMATION:
APPLICANT: Leo G.J. FRENKEN
APPLICANT: Cornelis P.E. van der LOGT
APPLICANT: Vin-Min TEH
APPLICANT: Martine E. VERHOEYEN
APPLICANT: Joy E. WILKINSON
APPLICANT: Stephen A. JOBLING
TITLE OF INVENTION: Production of Antibodies
FILE REFERENCE: PAK/060113/0275850 - T7060C
CURRENT APPLICATION NUMBER: US/09/737,476B
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: EP 99310188.0
PRIOR FILING DATE: 1999-12-17
NUMBER OF SEQ ID NOS: 67
SOFTWARE: MS Word
SEQ ID NO 56
LENGTH: 1172
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: hinge-HCV33-CH2-CH3-SEKDEL
NAME/KEY: CDS
LOCATION: (3) .. (1154)
US-09-737-476B-56
Alignment Scores:
Pred. No.: 0.347
Score: 76.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 90.48%
DB: 10
Gaps: 0
US-09-858-332g-15 (1-12) x US-09-737-476B-56 (1-1172)
QY 2 AsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 406 AACCAACCAACCAACCAACCAACCAACCAAC 438


```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-536

Alignment Scores:
Pred. No.:          6.56           Length:      5826
Score:              72.00         Matches:       10
Percent Similarity: 90.91%        Conservative: 0
Best Local Similarity: 90.91%     Mismatches:   1
Query Match:        85.71%        Indels:       0
DB:                 15           Gaps:         0

US-09-858-332G-15 (1-12) x US-10-311-455-536 (1-5826)
Oy      1 HisAsnHisAsnHisAsnHisAsnHis 11
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      335 CATACCACCAACCATTAACCCATTAACCATTAACAC 303

RESULT 42
US-10-311-455-1603/c
; Sequence 1603, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIERBENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Deter-
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311.455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 1603
LENGTH: 6236
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1603

Alignment Scores:
Pred. No.:          7.05           Length:      6236
Score:              72.00         Matches:       10
Percent Similarity: 90.91%        Conservative: 0
Best Local Similarity: 90.91%     Mismatches:   1
Query Match:        85.71%        Indels:       0
DB:                 15           Gaps:         0

US-09-858-332G-15 (1-12) x US-10-311-455-1603 (1-6236)
Oy      1 HisAsnHisAsnHisAsnHisAsnHis 11
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      4241 CATACCACCAACCATTAACCCATTAACCATTAACAC 4209

RESULT 43
US-10-312-841-2/c
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MR
FILE REFERENCE: E01/1208/WO
CURRENT APPLICATION NUMBER: US/10/312.841
CURRENT FILING DATE: 2002-12-30
NUMBER OF SEQ ID NOS: 2
SEQ ID NO 2
```

Alignment Scores:	
Pred. No.:	6.61e+03
Percent Similarity:	72.00
Best Local Similarity:	90.91%
Query Match:	85.71%
DB:	16
Length:	367377B
Matches:	10
Mismatches:	1
Indels:	0
Gaps:	0

QY 1 HisASnHtSASnHtSASnHtSASnHtSASnHtS 11
| | | | | | | | | | | | | | | | | | | | | |
Db 244224 CATACCACAAACCATATCCCATTAACCATTAACAC 244197

Alignment Scores:

5-09-858-332G-15 (1-12) X US-10-027-632-174961 (1-3186778)

```
1 116A6NH16A6NH16A6NH16A6NH16A6NH16 11
:::|||||
```

US-10-027-632-174961

Sequence 174961, Application US/10027632
Publication No. US20030204075A9
GENERAL TYPE OF INVENTION

CURRENT FILING DATE: 2002-04-30

TYPE: DNA

Alignment Scores:

S-09-858-332G-15 (1-12) X US-10-027-632-174961 (1-3186778)

5-09-817-310-1/c

APPLICANT: Stewart,

CURRENT FILING DATE: 2001-03-20

TYPE: DNA

ORGANISM: *Drosophila melanogaster*

```
US-09-817-310-1
Alignment Scores:
Pred. No.: 4.99      Length: 2556
Score: 70.00      Matches: 10
Percent Similarity: 91.67%      Conservative: 1
Best Local Similarity: 83.33%      Mismatches: 1
Query Match: 83.33%      Indels: 0
DB: 9      Gaps: 0

US-09-858-332g-15 (1-12) x US-09-817-310-1 (1-2556)
QY 1 HisasnHisasnHisasnHisasnHisasn 12
Db 2325 CATATCATATCATATCATATCATATCATATCATATCATAT 2230

RESULT 47
US-10-451-467A-641
Sequence 641, Application US/10451467A
GENERAL INFORMATION:
APPLICANT: CONTRERAS, ROLAND HENRI
APPLICANT: BERNHARDT, INES
APPLICANT: LUTTEN, WALTER HERMAN MARIA LOUIS
APPLICANT: REEKMAN, RIEKA JOSEPHINA
TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN
FILE REFERENCE: JAB-1667
CURRENT APPLICATION NUMBER: US/10/451,467A
CURRENT FILING DATE: 2003-06-19
PRIOR APPLICATION NUMBER: EP 00870318.3
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: EP 01870002.1
PRIOR FILING DATE: 2001-01-04
PRIOR APPLICATION NUMBER: EP 01870003.9
PRIOR FILING DATE: 2001-01-09
NUMBER OF SEQ ID NOS: 732
SOFTWARE: PatentIn version 3.1
SEQ ID NO 641
LENGTH: 24076
TYPE: DNA
ORGANISM: Candida albicans
US-10-451-467A-641

Alignment Scores:
Pred. No.: 55.3      Length: 24076
Score: 70.00      Matches: 10
Percent Similarity: 100.00%      Conservative: 2
Best Local Similarity: 83.33%      Mismatches: 0
Query Match: 83.33%      Indels: 0
DB: 19      Gaps: 0

US-09-858-332g-15 (1-12) x US-10-451-467A-641 (1-24076)
QY 1 HisasnHisasnHisasnHisasnHisasn 12
Db 22385 AACACCAACCAACCAACCAACCAACCAACCAACCAATAC 22420

RESULT 48
US-10-027-632-179264/c
Sequence 179264, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
```

```
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 179264
LENGTH: 1223197
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(1223197)
OTHER INFORMATION: n = A,T,C or G
US-10-027-632-179264

Alignment Scores:
Pred. No.: 3.74e+03      Length: 1223197
Score: 70.00      Matches: 10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 83.33%      Indels: 0
DB: 13      Gaps: 0

US-09-858-332g-15 (1-12) x US-10-027-632-179264 (1-1223197)
QY 2 AasnHisasnHisasnHisasnHisasn 11
Db 147154 AACCATACCATACCATACCATACCATACCAT 147125

RESULT 49
US-10-027-632-179264/c
Sequence 179264, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 179264
LENGTH: 1223197
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(1223197)
OTHER INFORMATION: n = A,T,C or G
US-10-027-632-179264
```

Alignment Scores:

Pred. No.: 3.74e+03 Length: 1223197
Score: 70.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 83.33% Indels: 0
DB: 17 Gaps: 0

US-09-858-332G-15 (1-12) x US-10-027-632-179264 (1-1223197)

OY 2 AsnHisAsnHisAsnHisAsnHisAsnHis 11

Db 147154 AACCATACCATACCATACCATACCATACCAT 147125

RESULT 50

/ Sequence 107879, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 107879
/ LENGTH: 550
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_68430C.1
US-10-424-599-107879

Alignment Scores:
Pred. No.: 1.3 Length: 550
Score: 69.00 Matches: 9
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 0
Query Match: 82.14% Indels: 0
DB: 18 Gaps: 0

US-09-858-332G-15 (1-12) x US-10-424-599-107879 (1-550)

OY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHis 12

Db 262 CATCACCATACCATACCATACCATACCATACCAT 297

RESULT 51

/ Sequence 57663, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 57663
/ LENGTH: 676
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_23081C.1
US-10-424-599-57663

Alignment Scores:

Pred. No.: 1.62 Length: 676
Score: 69.00 Matches: 9
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 0
Query Match: 82.14% Indels: 0
DB: 18 Gaps: 0

US-09-858-332G-15 (1-12) x US-10-424-599-57663 (1-676)

OY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHis 12

Db 205 CATACCATACCATACCATACCATACCATACCAT 240

RESULT 52

/ Sequence 50046, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J
/ APPLICANT: Kovalic, David K
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 50046
/ LENGTH: 1385
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_52571C.1
US-10-437-963-50046

Alignment Scores:
Pred. No.: 3.51 Length: 1385
Score: 69.00 Matches: 9
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 0
Query Match: 82.14% Indels: 0
DB: 19 Gaps: 0

US-09-858-332G-15 (1-12) x US-10-437-963-50046 (1-1385)

OY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHis 12

Db 574 CACAGCCACAGCCACAGCCACAGCCACAGCCACAGC 609

RESULT 53

/ Sequence 42398, Application US/11097143
/ Publication No. US20050208558A1
/ GENERAL INFORMATION:
/ APPLICANT: Venier, J. Craig
/ APPLICANT: et al
/ TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
/ FILE REFERENCE: C1000728
/ CURRENT APPLICATION NUMBER: US/11/097,143
/ PRIOR FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: 60/157,832
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/160,191
/ PRIOR FILING DATE: 1999-10-19

```
/ PRIOR APPLICATION NUMBER: 60/161,932
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: 60/164,769
/ PRIOR FILING DATE: 1999-11-12
/ PRIOR APPLICATION NUMBER: 60/173,383
/ PRIOR FILING DATE: 1999-12-28
/ PRIOR APPLICATION NUMBER: 60/175,693
/ PRIOR FILING DATE: 2000-01-12
/ PRIOR APPLICATION NUMBER: 60/184,831
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/191,637
/ PRIOR FILING DATE: 2000-03-23
/ NUMBER OF SEQ ID NOS: 43008
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 42398
/ LENGTH: 3855
/ TYPE: DNA
/ ORGANISM: DROSOPHILA
US-11-097-143-42398
```

```
Alignment Scores:
Pred. No.: 10.5      Length: 3855
Score: 69.00      Matches: 9
Percent Similarity: 100.00%      Conservative: 3
Best Local Similarity: 75.00%      Mismatches: 0
Query Match: 82.14%      Indels: 0
DB: 24      Gaps: 0
```

US-09-858-332g-15 (1-12) x US-11-097-143-42398 (1-3855)

```
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 1487 CATCATCATATATCATATAGTACAGCCACCAACCAAC 3522
```

```
RESULT 54
US-11-097-143-42397/c
/ Sequence 42397, Application US/11097143
/ Publication No. US20050208558A1
/ GENERAL INFORMATION:
/ APPLICANT: Venter, J. Craig
/ TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
/ TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
/ FILE REFERENCE: CL000728
/ CURRENT APPLICATION NUMBER: US/11/097,143
/ PRIOR FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: 60/157,832
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/160,191
/ PRIOR FILING DATE: 1999-10-19
/ PRIOR APPLICATION NUMBER: 60/161,932
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: 60/164,769
/ PRIOR FILING DATE: 1999-11-12
/ PRIOR APPLICATION NUMBER: 60/173,383
/ PRIOR FILING DATE: 1999-12-28
/ PRIOR APPLICATION NUMBER: 60/175,693
/ PRIOR FILING DATE: 2000-01-12
/ PRIOR APPLICATION NUMBER: 60/184,831
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/191,637
/ PRIOR FILING DATE: 2000-03-23
/ NUMBER OF SEQ ID NOS: 43008
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 42397
/ LENGTH: 10907
/ TYPE: DNA
/ ORGANISM: DROSOPHILA
US-11-097-143-42397
```

Alignment Scores: 32.1 Length: 10907

```
Score: 69.00      Matches: 9
Percent Similarity: 100.00%      Conservative: 3
Best Local Similarity: 75.00%      Mismatches: 0
Query Match: 82.14%      Indels: 0
DB: 24      Gaps: 0
```

US-09-858-332g-15 (1-12) x US-11-097-143-42397 (1-10907)

```
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 1369 CATCATCATATATCATATAGTACAGCCACCAACCAAC 1334
```

```
RESULT 55
US-11-097-143-21194
/ Sequence 21194, Application US/11097143
/ Publication No. US20050208558A1
/ GENERAL INFORMATION:
/ APPLICANT: Venter, J. Craig
/ TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
/ TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
/ FILE REFERENCE: CL000728
/ CURRENT APPLICATION NUMBER: US/11/097,143
/ PRIOR FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: 60/157,832
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/160,191
/ PRIOR FILING DATE: 1999-10-19
/ PRIOR APPLICATION NUMBER: 60/161,932
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: 60/164,769
/ PRIOR FILING DATE: 1999-11-12
/ PRIOR APPLICATION NUMBER: 60/173,383
/ PRIOR FILING DATE: 1999-12-28
/ PRIOR APPLICATION NUMBER: 60/175,693
/ PRIOR FILING DATE: 2000-01-12
/ PRIOR APPLICATION NUMBER: 60/184,831
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/191,637
/ PRIOR FILING DATE: 2000-03-23
/ NUMBER OF SEQ ID NOS: 43008
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 21194
/ LENGTH: 2094
/ TYPE: DNA
/ ORGANISM: DROSOPHILA
US-11-097-143-21194
```

```
Alignment Scores:
Pred. No.: 6.36      Length: 2094
Score: 68.50      Matches: 11
Percent Similarity: 92.31%      Conservative: 1
Best Local Similarity: 84.62%      Mismatches: 0
Query Match: 81.55%      Indels: 1
DB: 24      Gaps: 1
```

US-09-858-332g-15 (1-12) x US-11-097-143-21194 (1-2094)

```
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 1864 CACAACCATATACACCAACCAACCAACCAACCACTCC 1902
```

```
RESULT 56
US-11-097-143-21193/c
/ Sequence 21193, Application US/11097143
/ Publication No. US20050208558A1
/ GENERAL INFORMATION:
/ APPLICANT: Venter, J. Craig
/ TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
/ TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
/ FILE REFERENCE: CL000728
/ CURRENT APPLICATION NUMBER: US/11/097,143
/ PRIOR FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: 60/157,832
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/160,191
/ PRIOR FILING DATE: 1999-10-19
/ PRIOR APPLICATION NUMBER: 60/161,932
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: 60/164,769
/ PRIOR FILING DATE: 1999-11-12
/ PRIOR APPLICATION NUMBER: 60/173,383
/ PRIOR FILING DATE: 1999-12-28
/ PRIOR APPLICATION NUMBER: 60/175,693
/ PRIOR FILING DATE: 2000-01-12
/ PRIOR APPLICATION NUMBER: 60/184,831
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/191,637
/ PRIOR FILING DATE: 2000-03-23
/ NUMBER OF SEQ ID NOS: 43008
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 21193
/ LENGTH: 10907
/ TYPE: DNA
/ ORGANISM: DROSOPHILA
US-11-097-143-21193
```

Alignment Scores: 32.1 Length: 10907

```

; LENGTH: 4779
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-24139

Alignment Scores:

Pred. No.:      24.4      Length:      4779
Score:          67.00     Matches:      9
Percent Similarity: 83.33% Conservative: 1
Best Local Similarity: 75.00% Mismatches: 2
Query Match:      79.76% Indels:      0
DB:               24      Gaps:      0

US-09-858-332G-15 (1-12) x US-11-097-143-24139 (1-4779)

QY      1  HisAnH1sAenH1sAenH1sAenH1sAenH1sAen 12
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2475 CACACACCAAAACGACTCACACCAACCAACACACAC 2510

RESULT 58
US-10-424-599-96040
; Sequence 96040, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 96040
; LENGTH: 331
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57735C.1
US-10-424-599-96040

Alignment Scores:

Pred. No.:      1.89      Length:      331
Score:          66.00     Matches:      9
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 81.82% Mismatches: 0
Query Match:      78.57% Indels:      0
DB:               18      Gaps:      0

US-09-858-332G-15 (1-12) x US-10-424-599-96040 (1-331)

QY      2  AsnH1sAenH1sAenH1sAenH1sAenH1sAen 12
      ::|||::|||::|||::|||::|||::|||::|||
Db      40 AGTCACAGTCACATCATCAACCAACCAACCAAC 72

RESULT 59
US-10-425-115-19761
; Sequence 19761, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222) B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 19761
; LENGTH: 449

```

```

; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(449)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_118026C.1
US-10-425-115-19761

Alignment Scores:
Pred. No.: 2.62 Length: 449
Score: 66.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 78.57% Indels: 0
DB: 20 Gaps: 0

US-09-858-332g-15 (1-12) x US-10-425-115-19761 (1-449)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHis 11
Db 1 CACACCAACCAACCAACGCGCCACGCGCCACATCAC 33

RESULT 60
US-10-424-599-97126
; Sequence 97126, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 97126
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_58717C.1
US-10-424-599-97126

Alignment Scores:
Pred. No.: 8.42 Length: 1335
Score: 66.00 Matches: 9
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 81.82% Mismatches: 0
Query Match: 78.57% Indels: 0
DB: 18 Gaps: 0

US-09-858-332g-15 (1-12) x US-10-424-599-97126 (1-1335)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHis 11
Db 54 TACACCAACCAACCAACCACTACCAACCAACAC 86

RESULT 61
US-10-374-780A-996
; Sequence 996, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
```

```

; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Brown, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 996
; LENGTH: 1338
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Predicted polypeptide sequence is orthologous to G636
US-10-374-780A-996

Alignment Scores:
Pred. No.: 8.44 Length: 1338
Score: 66.00 Matches: 9
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 81.82% Mismatches: 0
Query Match: 78.57% Indels: 0
DB: 17 Gaps: 0

US-09-858-332g-15 (1-12) x US-10-374-780A-996 (1-1338)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHis 11
Db 54 TACACCAACCAACCAACCACTACCAACCAACAC 86

RESULT 62
US-10-412-699B-1275
; Sequence 1275, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Brown, Pierre E.
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
```

```

; APPLICANT: Creelman, Robert A.
; APPLICANT: Dubell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumamoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1275
; LENGTH: 1338
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-412-699B-1275

Alignment Scores:
Pred. No.:      8.44      Length:      1338
Score:          66.00     Matches:      9
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 81.82% Mismatches: 0
Query Match:    78.57%   Indels:      0
DB:             18       Gaps:        0

US-09-858-332G-15 (1-12) x US-10-412-699B-1275 (1-1338)
QY      1 HisAaNH1sAaNH1sAaNH1sAaNH1sAaNH1s 11
Db      54 TACAAACCAACCAACCACTACAAACCAACCAAC 86

RESULT 63
US-10-311-455-535/c
; Sequence 535, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 535
```

```

; LENGTH: 5826
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
; US-10-311-455-535

Alignment Scores:
Pred. No.:      40.9      Length:      5826
Score:          66.00     Matches:      9
Percent Similarity: 91.67% Conservative: 2
Best Local Similarity: 75.00% Mismatches: 1
Query Match:    78.57%   Indels:      0
DB:             15       Gaps:        0

US-09-858-332G-15 (1-12) x US-10-311-455-535 (1-5826)
QY      1 HisAaNH1sAaNH1sAaNH1sAaNH1sAaNH1s 12
Db      2029 TATAACCATTAACCAATTAACCAATTAACCAATTAAT 1994

RESULT 65
US-11-097-143-6454/c
; Sequence 6454, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.

Alignment Scores:
Pred. No.:      44       Length:      6236
Score:          66.00     Matches:      9
Percent Similarity: 91.67% Conservative: 2
Best Local Similarity: 75.00% Mismatches: 1
Query Match:    78.57%   Indels:      0
DB:             15       Gaps:        0

US-09-858-332G-15 (1-12) x US-10-311-455-1604 (1-6236)
QY      1 HisAaNH1sAaNH1sAaNH1sAaNH1sAaNH1s 12
Db      2029 TATAACCATTAACCAATTAACCAATTAACCAATTAAT 1994

RESULT 64
US-10-311-455-1604/c
; Sequence 1604, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1604
; LENGTH: 6236
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
; US-10-311-455-1604
```



```
FILE REFERENCE: CL000728
CURRENT APPLICATION NUMBER: US/11/097,143
CURRENT FILING DATE: 2005-04-04
PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/160,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/164,769
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/173,383
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: 60/175,693
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: 60/184,831
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637
PRIOR FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 43008
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6454
LENGTH: 28360
TYPE: DNA
ORGANISM: DROSOPHILA
US-11-097-143-6454
```

```
Alignment Scores:
Pred. No.: 223 Length: 28360
Score: 66.00 Matches: 10
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 2
Query Match: 78.57% Indels: 0
DB: 24 Gaps: 0
```

US-09-858-332G-15 (1-12) x US-11-097-143-6454 (1-28360)

```
QY 1 HisAenHisAenHisAenHisAenHisAenHisAen 12
Db 22869 CACAATCAGATCAGATCAGATCAGATCAGATCAGAT 22834
```

```
RESULT 66
US-10-312-841-1/c
Sequence 1, Application US/10312841
Publication No. US20030186277A1
GENERAL INFORMATION:
APPLICANT: Epigenomics AG
TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
FILE REFERENCE: E01/1208/WO
CURRENT APPLICATION NUMBER: US/10/312,841
CURRENT FILING DATE: 2002-12-30
NUMBER OF SEQ ID NOS: 2
SEQ ID NO 1
LENGTH: 3673778
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
NAME/KEY: unsure
LOCATION: (3294164)
US-10-312-841-1
```

```
Alignment Scores:
Pred. No.: 4,12e+04 Length: 3673778
Score: 66.00 Matches: 9
Percent Similarity: 91.67% Conservative: 2
Best Local Similarity: 75.00% Mismatches: 1
Query Match: 78.57% Indels: 0
DB: 16 Gaps: 0
```

US-09-858-332G-15 (1-12) x US-10-312-841-1 (1-3673778)

```
QY 1 HisAenHisAenHisAenHisAenHisAenHisAen 12
Db 3429588 TATAACCATTAACCATTAACCATTAACCATTAAT 3429553
RESULT 67
US-09-728-445-386
Sequence 386, Application US/09728445
Patent No. US20020102543A1
GENERAL INFORMATION:
APPLICANT: Friedrich, Glenn
APPLICANT: Zambrowicz, Brian
TITLE OF INVENTION: No. US20020102543A1 Mutated Mammalian Cells and
TITLE OF INVENTION: Animals
FILE REFERENCE: LEX-0102-USA
CURRENT APPLICATION NUMBER: US/09/728,445
CURRENT FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/168,358
PRIOR FILING DATE: 1999-12-01
NUMBER OF SEQ ID NOS: 891
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 386
LENGTH: 334
TYPE: DNA
ORGANISM: Mus musculus
US-09-728-445-386
```

```
Alignment Scores:
Pred. No.: 2.58 Length: 334
Score: 65.00 Matches: 9
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 0
Query Match: 77.38% Indels: 0
DB: 9 Gaps: 0
```

US-09-858-332G-15 (1-12) x US-09-728-445-386 (1-334)

```
QY 1 HisAenHisAenHisAenHisAenHisAenHisAen 12
Db 180 AACACACACACACACACACACACACACACACACAC 215
```

```
RESULT 68
US-10-964-549-386
Sequence 386, Application US/10964549
Publication No. US20050186677A1
GENERAL INFORMATION:
APPLICANT: Friedrich, Glenn
APPLICANT: Zambrowicz, Brian
APPLICANT: Sande, Arthur T.
TITLE OF INVENTION: Novel Mutated Mammalian Cells and Animals
FILE REFERENCE: LEX-0286-USA
CURRENT APPLICATION NUMBER: US/10/964,549
CURRENT FILING DATE: 2004-10-13
PRIOR APPLICATION NUMBER: US/09/750,456
PRIOR FILING DATE: 2000-12-28
PRIOR APPLICATION NUMBER: US 09/728,445
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/168,358
PRIOR FILING DATE: 1999-12-01
NUMBER OF SEQ ID NOS: 891
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 386
LENGTH: 334
TYPE: DNA
ORGANISM: Mus musculus
US-10-964-549-386
```

```
Alignment Scores:
Pred. No.: 2.58 Length: 334
Score: 65.00 Matches: 9
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 0
Query Match: 77.38% Indels: 0
```



```
/ Publication No. US20050208558A1
/ GENERAL INFORMATION:
/ APPLICANT: Venter, J. Craig
/ APPLICANT: et al
/ TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
/ TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
/ FILE REFERENCE: DROSOPHILA GENES.
/ FILE REFERENCE: CL000728
/ CURRENT APPLICATION NUMBER: US/11/097,143
/ CURRENT FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: 60/157,832
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/160,191
/ PRIOR FILING DATE: 1999-10-19
/ PRIOR APPLICATION NUMBER: 60/161,932
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: 60/164,769
/ PRIOR FILING DATE: 1999-11-12
/ PRIOR APPLICATION NUMBER: 60/173,383
/ PRIOR FILING DATE: 1999-12-28
/ PRIOR APPLICATION NUMBER: 60/175,693
/ PRIOR FILING DATE: 2000-01-12
/ PRIOR APPLICATION NUMBER: 60/184,831
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/191,637
/ PRIOR FILING DATE: 2000-03-23
/ NUMBER OF SEQ ID NOS: 43008
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 18322
/ LENGTH: 13033
/ TYPE: DNA
/ ORGANISM: DROSOPHILA
/ US-11-097-143-18322

Alignment Scores:
Pred. No.: 132 Length: 13033
Score: 65.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 77.38% Indels: 0
DB: 24 Gaps: 0

US-09-858-332g-15 (1-12) x US-11-097-143-18322 (1-13033)
Qy 1 H1sAsnH1sAsnH1sAsnH1sAsnH1sAsnH1s 11
Db 8114 CACACCACCAACCAACCAACCAACCAACCAAC 8082

RESULT 73
US-10-398-221-9
/ Sequence 9, Application US/10398221
/ Publication No. US20040018514A1
/ GENERAL INFORMATION:
/ APPLICANT: KUNST, Frederik
/ APPLICANT: GLASER, Philippe
/ TITLE OF INVENTION: Listeria innocua, genome and applications
/ FILE REFERENCE: 344 702 - US
/ CURRENT APPLICATION NUMBER: US/10/398,221
/ CURRENT FILING DATE: 2003-03-27
/ PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
/ PRIOR FILING DATE: 2001-10-04
/ PRIOR APPLICATION NUMBER: FR 00/12 697
/ PRIOR FILING DATE: 2000-10-04
/ NUMBER OF SEQ ID NOS: 4025
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 9
/ LENGTH: 684707
/ TYPE: DNA
/ ORGANISM: Listeria innocua
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(end)
/ OTHER INFORMATION: n can be any nucleotide: a, g, c or t/u
```

```
US-10-398-221-9
Alignment Scores:
Pred. No.: 9,22e+03 Length: 684707
Score: 65.00 Matches: 9
Percent Similarity: 83.33% Conservative: 1
Best Local Similarity: 75.00% Mismatches: 2
Query Match: 77.38% Indels: 0
DB: 17 Gaps: 0

US-09-858-332g-15 (1-12) x US-10-398-221-9 (1-684707)
Qy 1 H1sAsnH1sAsnH1sAsnH1sAsnH1sAsnH1s 12
Db 542149 CATATCATCATCATCATCATGTGTCAACCATTAAC 542184

RESULT 74
US-10-398-221-2058/c
/ Sequence 2058, Application US/10398221
/ Publication No. US20040018514A1
/ GENERAL INFORMATION:
/ APPLICANT: KUNST, Frederik
/ APPLICANT: GLASER, Philippe
/ TITLE OF INVENTION: Listeria innocua, genome and applications
/ FILE REFERENCE: 344 702 - US
/ CURRENT APPLICATION NUMBER: US/10/398,221
/ CURRENT FILING DATE: 2003-03-27
/ PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
/ PRIOR FILING DATE: 2001-10-04
/ PRIOR APPLICATION NUMBER: FR 00/12 697
/ PRIOR FILING DATE: 2000-10-04
/ NUMBER OF SEQ ID NOS: 4025
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 2058
/ LENGTH: 3011208
/ TYPE: DNA
/ ORGANISM: Listeria innocua
/ US-10-398-221-2058

Alignment Scores:
Pred. No.: 4.52e+04 Length: 3011208
Score: 65.00 Matches: 9
Percent Similarity: 83.33% Conservative: 1
Best Local Similarity: 75.00% Mismatches: 2
Query Match: 77.38% Indels: 0
DB: 17 Gaps: 0

US-09-858-332g-15 (1-12) x US-10-398-221-2058 (1-3011208)
Qy 1 H1sAsnH1sAsnH1sAsnH1sAsnH1sAsnH1s 12
Db 2738663 CATATCATCATCATCATCATGTGTCAACCATTAAC 2738628

RESULT 75
US-10-032-585-1299/c
/ Sequence 1299, Application US/10032585
/ Publication No. US20030180953A1
/ GENERAL INFORMATION:
/ APPLICANT: Terry, Roemer D.
/ APPLICANT: Bo, Jians
/ APPLICANT: Charles, Boone
/ APPLICANT: Howard, Bussey
/ TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
/ FILE REFERENCE: 10182-005-999
/ CURRENT APPLICATION NUMBER: US/10/032,585
/ CURRENT FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 8000
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1299
/ LENGTH: 65
/ TYPE: DNA
/ ORGANISM: Candida albicans
/ US-10-032-585-1299
```

Alignment Scores:

Pred. No.:	0.606	Length:	65
Score:	64.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	4
Best Local Similarity:	66.67%	Mismatches:	0
Query Match:	76.19%	Indels:	0
DB:	16	Gaps:	0

US-09-858-332g-15 (1-12) x US-10-032-585-1299 (1-65)

Qy 1 H1s8nhi8a8nhi8a8nhi8a8nhi8a8nhi8a8n 12

Db 64 CATTACCATTAACACACACACACACACACACACACACACAC 29

RESULT 76

US-09-864-761-19229
Sequence 19229, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 19229
LENGTH: 148
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC010876.1

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
OTHER INFORMATION: EST HUMAN HIT: AA021158.1, EVALUATE 2.00e+00
OTHER INFORMATION: NT HIT: AL161565.2, EVALUATE 2.80e-01
US-09-864-761-19229

Alignment Scores:

Pred. No.:	1.46	Length:	148
Score:	64.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	4
Best Local Similarity:	66.67%	Mismatches:	0
Query Match:	76.19%	Indels:	0
DB:	9	Gaps:	0

US-09-858-332g-15 (1-12) x US-09-864-761-19229 (1-148)

Qy 1 H1s8nhi8a8nhi8a8nhi8a8nhi8a8nhi8a8n 12

Db 11 CACACACACACACACACACACACACACACACACACACAC 46

RESULT 77

US-09-864-761-2500
Sequence 2500, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687

```

; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2500
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010876.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
US-09-864-761-2500

Alignment Scores:
Pred. No.: 4.04 Length: 381
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 9 Gaps: 0

US-09-858-332g-15 (1-12) x US-09-864-761-2500 (1-381)
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 274 CACAACCAACACCAACCAACCAACCAACCAACCAACCAAC 309

RESULT 78
US-10-424-599-51775/c
; Sequence 51775, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21 (53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 51775
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_17767C.1
US-10-424-599-51775

Alignment Scores:
Pred. No.: 4.24 Length: 399
Score: 64.00 Matches: 9
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 90.00% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 18 Gaps: 0

US-09-858-332g-15 (1-12) x US-10-424-599-51775 (1-399)
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 10
```

```

Db 368 TACAATCATATACCATATATATATACCATATAT 339
RESULT 79:
US-10-674-124A-3759/c
; Sequence 3759, Application US/10674124A
; Publication No. US2004019797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hideotoshi
; APPLICANT: TAMURA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; TITLE OF INVENTION: GENETIC POLYMORPHISM MARKERS
; FILE REFERENCE: ORIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; PRIOR FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 3759
; LENGTH: 413
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: chr2.fa.07ftr.169288578
; FEATURE:
; OTHER INFORMATION: Located on chromosome 2
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; OTHER INFORMATION: sequence : 163393275
; FEATURE:
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 126010
US-10-674-124A-3759

Alignment Scores:
Pred. No.: 4.4 Length: 413
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 19 Gaps: 0

US-09-858-332g-15 (1-12) x US-10-674-124A-3759 (1-413)
QY 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 258 CACAACCAACCAACCAACCAACCAACCAACCAACCAAC 223

RESULT 80
US-09-770-444-818
; Sequence 818, Application US/09770444
; Patent No. US20020023280A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Mathew, Abraham V.
; APPLICANT: Ledford, Brooke L.
```

```

1  PRIOR FILING DATE: 2001-01-30
2  PRIOR APPLICATION NUMBER: PCT/US01/00665
3  PRIOR FILING DATE: 2001-01-30
4  PRIOR APPLICATION NUMBER: PCT/US01/00668
5  PRIOR FILING DATE: 2001-01-30
6  PRIOR APPLICATION NUMBER: PCT/US01/00663
7  PRIOR FILING DATE: 2001-01-30
8  PRIOR APPLICATION NUMBER: PCT/US01/00662
9  PRIOR FILING DATE: 2001-01-30
10 PRIOR APPLICATION NUMBER: PCT/US01/00661
11 PRIOR FILING DATE: 2001-01-30
12 PRIOR APPLICATION NUMBER: PCT/US01/00670
13 PRIOR FILING DATE: 2001-01-30
14 PRIOR APPLICATION NUMBER: US 60/234,687
15 PRIOR FILING DATE: 2000-09-21
16 PRIOR APPLICATION NUMBER: US 09/608,408
17 PRIOR FILING DATE: 2000-06-30
18 PRIOR APPLICATION NUMBER: US 09/774,203
19 PRIOR FILING DATE: 2001-01-29
20 NUMBER OF SEQ ID NOS: 49117
21 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
22 SEQ ID NO 768
23 LENGTH: 479
24 TYPE: DNA
25 ORGANISM: Homo sapiens
26 FEATURE:

```

1	OTHER INFORMATION:	EXPRESSED IN HELIA, SIGNAL = 12
2	OTHER INFORMATION:	MAP TO 28.851.17
3	OTHER INFORMATION:	EXPRESSED IN ADULT LIVER, SIGNAL = 11
4	OTHER INFORMATION:	EXPRESSED IN BRAIN, SIGNAL = 12
5	OTHER INFORMATION:	EXPRESSED IN HEART, SIGNAL = 7
6	OTHER INFORMATION:	EXPRESSED IN PLACENTA, SIGNAL = 12
7	OTHER INFORMATION:	EXPRESSED IN BT474, SIGNAL = 11
8	OTHER INFORMATION:	EXPRESSED IN BONE MARROW, SIGNAL = 13
9	OTHER INFORMATION:	EXPRESSED IN HEL100, SIGNAL = 11
10	OTHER INFORMATION:	EXPRESSED IN FETAL LIVER, SIGNAL = 11

US-09-864-761-768

Alignment Scores:

Score:	5.16	Length:	479
Percent Similarity:	64.00	Matches:	8
Best Local Similarity:	100.00	Conservative:	4
Query Match:	66.67	Mismatches:	0
	76.19	Indels:	0

US-09-858-332G-15 (1-12) x US-09-864-761-768 (1-479)

Qy 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
|||:|||||:|||||:|||||:|||||:|||||
Db 457 CACCACCATACCAACCAACCAACCAACCATAC 422

RESULT 82
US-09-770-961-127/C
; Sequence 127, Application US/09770961
; Publication No. US20030115539A1
GENERAL INFORMATION:
APPLICANT: Corlath, John
APPLICANT: Ma, Yong-Qiang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Kameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Mathew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Woessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Kitcher, MaJa

APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurlan, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
FILE REFERENCE: 2026 (PARA-015PRV)
CURRENT APPLICATION NUMBER: US/09/770,961
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/178,466
PRIOR FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 127
LENGTH: 515
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-770-961-127

Alignment Scores:
Pred. No.: 5.58 Length: 515
Score: 64.00 Matches: 9
Percent Similarity: 91.67% Conservative: 2
Best Local Similarity: 75.00% Mismatches: 1
Query Match: 76.19% Indels: 0
DB: 10 Gaps: 0

US-09-858-332g-15 (1-12) x US-09-770-961-127 (1-515)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 431 AACACCAACGACATACCAAAACCAACCAACCAAT 396

RESULT 83
US-10-029-386-3710
Sequence 3710, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: AEOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 3710
LENGTH: 569
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR22 180.0
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.78
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
OTHER INFORMATION: SWISSPROT HIT: P02583, EVALUAE 9.40e+00
OTHER INFORMATION: EST HUMAN HIT: BF961392.1, EVALUAE 1.00e-103
OTHER INFORMATION: NT HIT: J05096.1, EVALUAE 2.10e+00
US-10-029-386-3710

Alignment Scores:
Pred. No.: 6.21 Length: 569
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 16 Gaps: 0

US-09-858-332g-15 (1-12) x US-10-029-386-3710 (1-569)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 161 AACACCAACGACATACCAAAACCAACCAACCAAT 127

Db 37 CACCAACATCACCATATACCAACCAACATATACCAAT 72

RESULT 84

US-10-767-701-26149/C
Sequence 26149, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 26149
LENGTH: 609
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: 30970996
US-10-767-701-26149

Alignment Scores:
Pred. No.: 6.68 Length: 609
Score: 64.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 19 Gaps: 0

US-09-858-332g-15 (1-12) x US-10-767-701-26149 (1-609)

QY 1 HisAsnHisAsnHisAsnHisAsnHis 9
Db 492 CATTAACCATTAACCATTAACCATTAACCATTAAC 466

RESULT 85
US-10-424-599-109460/C
Sequence 109460, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Kovalic, David K
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 109460
LENGTH: 629
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(629)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: PAT_MRT3847_69858C.1
US-10-424-599-109460

Alignment Scores:
Pred. No.: 6.91 Length: 629
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 18 Gaps: 0

OTHER INFORMATION: EXPRESSED IN FETAL

Length: 15.8

Pred. No.:	15.8	Length:	1356
Score:	64.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	4


```

; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22081
; LENGTH: 12396
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-22081
```

```

Alignment Scores:
Pred. No.: 169 Length: 12396
Score: 64.00 Matches: 9
Percent Similarity: 83.33% Conservative: 1
Best Local Similarity: 75.00% Mismatches: 2
Query Match: 76.19% Indels: 0
DB: 24 Gaps: 0
```

US-09-858-332g-15 (1-12) x US-11-097-143-22081 (1-12396)

```
Oy 1 Hisaenhisaenhisaenhisaenhisaenhisaen 12
Db 11338 CATTAACATTCATTCATTCATTCATTCATTCATTAAT 11303
```

```

RESULT 92
US-11-097-143-42484
; Sequence 42484, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CLO00728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42484
; LENGTH: 21379
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-42484
```

```

Alignment Scores:
Pred. No.: 304 Length: 21379
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 24 Gaps: 0
```

US-09-858-332g-15 (1-12) x US-11-097-143-42484 (1-21379)

```
Oy 1 Hisaenhisaenhisaenhisaenhisaenhisaen 12
Db 16826 CATTAACATTCATTCATTCATTCATTCATTCATTCATTC 16861
```

RESULT 93

```

US-11-097-143-42451
; Sequence 42451, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CLO00728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42451
; LENGTH: 21409
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-42451
```

```

Alignment Scores:
Pred. No.: 304 Length: 21409
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 24 Gaps: 0
```

US-09-858-332g-15 (1-12) x US-11-097-143-42451 (1-21409)

```
Oy 1 Hisaenhisaenhisaenhisaenhisaenhisaen 12
Db 16856 CATTAACATTCATTCATTCATTCATTCATTCATTCATTC 16891
```

RESULT 94

```

US-11-097-143-2977
; Sequence 2977, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
```

```
/ TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
/ TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
/ TITLE OF INVENTION: DROSOPHILA GENES.
/ FILE REFERENCE: CL000728
/ CURRENT APPLICATION NUMBER: US/11/097,143
/ CURRENT FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: 60/157,832
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/160,191
/ PRIOR FILING DATE: 1999-10-19
/ PRIOR APPLICATION NUMBER: 60/161,932
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: 60/164,769
/ PRIOR FILING DATE: 1999-11-12
/ PRIOR APPLICATION NUMBER: 60/173,383
/ PRIOR FILING DATE: 1999-12-28
/ PRIOR APPLICATION NUMBER: 60/175,693
/ PRIOR FILING DATE: 2000-01-12
/ PRIOR APPLICATION NUMBER: 60/184,831
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/191,637
/ PRIOR FILING DATE: 2000-03-23
/ NUMBER OF SEQ ID NOS: 43008
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2977
/ LENGTH: 21449
/ TYPE: DNA
/ ORGANISM: DROSOPHILA
US-11-097-143-2977

Alignment Scores:
Pred. No.: 305 Length: 21449
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 24 Gaps: 0

US-09-858-332g-15 (1-12) x US-11-097-143-2977 (1-21449)
Qy 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 16543 CACCACCAACCAACCAACCAACCAACCAACCAACCAAC 16578

RESULT 95
US-11-097-143-28753
/ Sequence 28753, Application US/11097143
/ Publication No. US20050208558A1
/ GENERAL INFORMATION:
/ APPLICANT: Venter, J. Craig
/ TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
/ TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
/ FILE REFERENCE: CL000728
/ CURRENT APPLICATION NUMBER: US/11/097,143
/ CURRENT FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: 60/157,832
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/160,191
/ PRIOR FILING DATE: 1999-10-19
/ PRIOR APPLICATION NUMBER: 60/161,932
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: 60/164,769
/ PRIOR FILING DATE: 1999-11-12
/ PRIOR APPLICATION NUMBER: 60/173,383
/ PRIOR FILING DATE: 1999-12-28
/ PRIOR APPLICATION NUMBER: 60/175,693
/ PRIOR FILING DATE: 2000-01-12
/ PRIOR APPLICATION NUMBER: 60/184,831
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/191,637
/ PRIOR FILING DATE: 2000-03-23
```

```
/ NUMBER OF SEQ ID NOS: 43008
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 28753
/ LENGTH: 21451
/ TYPE: DNA
/ ORGANISM: DROSOPHILA
US-11-097-143-28753

Alignment Scores:
Pred. No.: 305 Length: 21451
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 24 Gaps: 0

US-09-858-332g-15 (1-12) x US-11-097-143-28753 (1-21451)
Qy 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 16545 CACCACCAACCAACCAACCAACCAACCAACCAACCAAC 16580

RESULT 96
US-11-097-143-28756
/ Sequence 28756, Application US/11097143
/ Publication No. US20050208558A1
/ GENERAL INFORMATION:
/ APPLICANT: Venter, J. Craig
/ TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
/ TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
/ FILE REFERENCE: CL000728
/ CURRENT APPLICATION NUMBER: US/11/097,143
/ CURRENT FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: 60/157,832
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/160,191
/ PRIOR FILING DATE: 1999-10-19
/ PRIOR APPLICATION NUMBER: 60/161,932
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: 60/164,769
/ PRIOR FILING DATE: 1999-11-12
/ PRIOR APPLICATION NUMBER: 60/173,383
/ PRIOR FILING DATE: 1999-12-28
/ PRIOR APPLICATION NUMBER: 60/175,693
/ PRIOR FILING DATE: 2000-01-12
/ PRIOR APPLICATION NUMBER: 60/184,831
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/191,637
/ PRIOR FILING DATE: 2000-03-23
/ NUMBER OF SEQ ID NOS: 43008
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 28756
/ LENGTH: 21451
/ TYPE: DNA
/ ORGANISM: DROSOPHILA
US-11-097-143-28756

Alignment Scores:
Pred. No.: 305 Length: 21451
Score: 64.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 76.19% Indels: 0
DB: 24 Gaps: 0

US-09-858-332g-15 (1-12) x US-11-097-143-28756 (1-21451)
Qy 1 HisAsnHisAsnHisAsnHisAsnHisAsn 12
Db 16545 CACCACCAACCAACCAACCAACCAACCAACCAACCAAC 16580
```

Alignment Scores:	
Pred. No.:	579
Score:	64.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	76.19%
DB:	18
	Gaps: 0
	Indels: 0
	Mismatches: 0
	Conservative: 0
	Matches: 9
	Length: 39003

US-658-3326-15 (1-12) X US-10-672-787-21 (1-39003)

Oy 1 Hisasnhisasnhisasnhis 9
 |||||
 Db 3052 CACAATCATATCATATCACAACCAT 3078

```

RESULT 96
US-10-087-192-1459/c
; Sequence 1459, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Moritz, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1459
; LENGTH: 126192
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(126192)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1459

```

Alignment Scores:

File: NO.:	2.04e+03	Length:	126192
Score:	64.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	66.67%	Mismatches:	4
Query Match:	76.19%	Indels:	0
DB:	13	Gaps:	0

US-09-858-332G-15 (1-12) X US-10-087-192-1459 (1-126192)

Oy	1	HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn	12
Db	104003	CACCATCATTAACCATTAATCACCATCATCATCACCATCAT	103000

RESULT 99
US-10-240-425-1585

;; sequence 1585, Application US/10240425
;; Publication No. US20040033502A1
;; GENERAL INFORMATION:
;; ADDITIONAL INFORMATION:

```

1  APPLICANT: Williams, Amanda
2  APPLICANT: Boland, Joseph F.
3  APPLICANT: Lord, Reginald V.
4  APPLICANT: Alvarez, Chris
5  APPLICANT: Wetzel, Jon C.
6  APPLICANT: Scherf, Iwe
7  APPLICANT: Vockley, Joseph G.
8  TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
9  FILE REFERENCE: 44921-5026
10 CURRENT APPLICATION NUMBER: 2002-09-10/240.425
11 CURRENT FILING DATE: 2002-09-30
12 PRIOR APPLICATION NUMBER: PCT/US01/09847
13 PRIOR FILING DATE: 2001-03-28
14 PRIOR APPLICATION NUMBER: US 60/193,446
15 PRIOR FILING DATE: 2000-03-31
16 NUMBER OF SEQ ID NOS: 1588
17 SOFTWARE: PatentIn Ver. 2.1
18 SEQ ID NO 1585
19 LENGTH: 135259
20 TYPE: DNA
21 ORGANISM: Homo sapiens
22 FEATURE:
23 OTHER INFORMATION: Genbank Accession No. US20040033502A1 Z83838
24 US-10-240-425-1585

```

Alignment Scores:

File: NO.:	2_2e+03	Length:	155253
Score:	64.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	4
Best Local Similarity:	66.67%	Mismatches:	0
Query Match:	76.19%	Indels:	0
DB:	18	Gaps:	0

US-10-240-425-1585 (1-12) X US-10-240-425-1585 (1-135259)

[illegible]

70048 CACCACCATATCACCACCATATCACCAT 70083

RESULT 100
US-10-026-188-3/c
; Sequence 3, Application US/10026188
; Publication No. US20020166645A1
; GENERAL INFORMATION:
IDENTIFICATION:

APPLICANT: Zhang, Yifeng

RESEARCHER: The Regents of the University of California
 TITLE OF INVENTION: Assays for Taste Receptor Cell Specificity
 TITLE OF INVENTION: Ion Channel
 FILE REFERENCE: 02307E-1149100S

CURRENT FILING DATE 2001-12-31

PRIOR FILING DATE: 2000-12-30

SOFTWARE: FASTSEN for Windows Version 1.0
NUMBER OF SEQ ID NOS: 8

LENGTH: 249487

ORGANISM: *Mus musculus*

FEATURE:

OTHER INFORMATION: mouse genomic region containing 1trpc5
US-10-026-188-3

Alignment Scores:

Pred. No.:	4.24e+03	Length:	249487
Score:	64.00	Matches:	9
Percent Similarity:	91.67%	Conservative:	2
Best Local Similarity:	75.00%	Mismatches:	1
Query Match:	76.19%	Indels:	0
DB:	13	Gaps:	0

US-09-858-332g-15 (1-12) x US-10-026-188-3 (1-249487)

QY 1 HisAsnHisAsnHisAsnHisAsnHisAsnHisAsn 12

DB 243511 CACACCCACACATTAATCATTAATCACAATTAATTAAC 243476

Search completed: October 4, 2005, 13:51:14
Job time : 1730 secs

THIS PAGE BLANK (USPTO)